



Child-Gender Preference Generalized Maximum Entropy Approach

Supanika Leurcharusmee[†], Jirakom Sirisrisakulchai^{‡,1},
Natthaphat Kingnetr[†], and Songsak Sriboonchitta[‡]

[†]Faculty of Economics, Chiang Mai University
Chiang Mai University, Chiang Mai 50200, Thailand
e-mail : supanika.econ.cmu@gmail.com (S. Leurcharusmee)
natthaphat.kingnetr@outlook.com (N. Kingnetr)

[‡]Center of Excellence in Econometrics, Faculty of Economics
Chiang Mai University, Chiang Mai 50200, Thailand
e-mail : sirisrisakulchai@hotmail.com (J. Sirisrisakulchai)
songsakecon@gmail.com (S. Sriboonchitta)

Abstract : This study measures child-gender preference in Thailand by examining the probability of having an additional child given genders of the previous children using the generalized maximum entropy (GME) approach. The GME method is robust to multicollinearity problem, which allows us to examine different child-gender preference pattern among families with different characteristics. The results show that girls are preferred in the families where the mothers have no high school degree. However, boys are preferred in the families where the mothers have a high school degree. Moreover, regardless of mothers' levels of education, the results show the evidence of mixed gender preference. Thai parents are less likely to have an additional child once they already have at least one boy and one girl.

Keywords : gender preference; generalized maximum entropy; Thailand.

2010 Mathematics Subject Classification : 62P20; 91B84.

This research was supported in part by Center of Excellence in Econometrics.
¹Corresponding author.

1 Introduction

Over the decades, social scientists have investigated child-gender preference and how it could affect fertility decisions and several other social issues. It has been studied that parents' marital and labor supply decisions can be affected by the genders of their children. There are several economic based hypotheses to explain such relationships involving different costs and health conditions of children of different genders [1]. However, several hypotheses were also developed under the assumption of the existence of child-gender bias. For example, in a family with a boy preference, a mother with a boy may have a higher negotiation power and is required to work less in the labor market [2]. Moreover, with boy preference, a family with boys may have higher marriage stability [3]. Therefore, the knowledge of child-gender preference in a society allows researchers to more clearly identify the reasoning toward the relationships between child gender and the household's marital and labor market outcomes.

Despite the importance of the topic, there has not been many studies in the case of Thailand. The most recent studies were more than a decades ago. The evidence may no longer represents the present situation due to changes in social and political structures as well as substantial economic development over the past ten years. Therefore, the purpose of this paper is to re-examine the existence of child gender preference across families in Thailand. Unlike previous studies in Thailand, we followed [4]'s approach to investigate the existence of child-gender preference from the parents' decision to have an additional child given the genders of the previous children. Specifically, we examined the decision to have the $(m + 1)^{st}$ children given the genders of the previous m children. The model can identify the boy preference, girl preference and mixed gender preference. If the parents have a boy preference, then the parents with a least one boy are less likely to have an additional child. If the parents have a mixed gender preference, then parents with at least one boy and one girl are less likely to have an additional child. Such model allows us to find the child-gender preference without the influence of the birth order and number of children.

For the empirical study, this study used the data from the 2015 Social Economic Survey (SES) collected by Thailand's National Statistical Office. With the limited number of observations for families with more than three children and potential multicollinearity problem of explanatory variables, this study adopted the [5]'s Generalize Maximum Entropy (GME) approach for a discrete choice model to estimate a series of decisions to have children. With the semi-parametric nature, the GME estimation is robust to distributional assumption of the errors. Moreover, numerous family characteristics can be taken into account in the study as the methodology is robust to multicollinearity problem. This study compares the traditional logit-based estimation with the entropy-based estimation and found that parameters estimated using the entropy-based method presented smaller standard errors, especially when the sample size is small.

Regarding the child-gender preference, the results show a complex structure of child-gender preference in Thailand. Specifically, in the families where the mothers

have education lower than high school, girls are preferred. In the families where the mothers have a high school degree, boys are preferred. However, regardless of mothers' levels of education, the results show evidence that Thai families have mixed gender preference. That is, parents are less likely to have an additional child once they already have at least one boy and one girl.

2 Literature Review

Child-gender preference has been studied in many countries, mostly in Asia where the preference for sons seems to be relatively higher than the rest of the world [6]. Based on the sex-ratios at birth (SRB) across 61 countries in Africa, Asia, and Latin America, [7] found that there exists son preference in Asia, especially in India and China, and Africa, while there is a little evidence to be seen in the case of Latin America. Even though previous studies suggest that families in Asian countries prefer to have sons over daughter, some recent studies show otherwise. [8] found that China, India and Vietnam exhibit strong son preference. However, based on the data from national survey over 1991 to 2012, the trend seems to be opposite in the case of South Korea [6]. They argued that social and economic development such as lower wage discrimination, greater gender equality, and higher income plays contribute to the current trend of non-bias in child-gender preference in South Korea.

In the case of Thailand, the research on child-gender preference has been rather limited and outdated. [9] were among the very first researchers to examine such preference. Relying on social-demographic survey and descriptive statistics, their results suggest little evidence for child-gender preference. In general, Thai families prefer to have children of both sexes. In addition, they found that the important factor affecting the decision on having additional children is the number of children that families have already had. Apart from the number of children, [10] suggest that urbanisation could also affect the reproductive behaviour of Thai women based on their two-round survey data. Later on, [11] re-examined gender preference for children across Thai families using the data from Thai national survey. Unlike previous studies, their results show that there is a difference in children-gender preference between men and women. they found that married women prefer to have children on both sexes, while men as well as families in the urban ethnic Chinese group show a strong preference for sons.

After two decades, [12] investigated the child-gender preference in South-East Asian countries using Coomb's scale as a measurement of child-gender preference. In the case of Thailand, their findings confirm what the previous studies have found. That is, Thai families have a mixed gender preference, which is that they prefer to have children of both genders. However, they did not investigate the determinants of child-gender preference due to the lack of data and appropriate methodology at the time. The findings was further confirmed by [13]'s study based on national surveys. The study found that the size of Thai families tend to limit at two children with a child for each gender.

Although these existing studies provide a useful foundation for the analysis of child-gender preference in Thailand, there are limitations in their methodologies since they only relied on survey data and descriptive statistics. This suggests that their findings may not truly reflect the current state of child-gender preferences across Thai families. Moreover, [12] mentioned in their study that, for future research, investigation on the determinants of child-gender preference should be done for policy maker to design appropriate solutions to deal with social issues involving the gender discrimination on children.

3 Model Formulation

In this study, we measure child-gender preference by estimating the probability of having an additional child given the genders of the previous children. That is, We formulate a mechanism for child gender preference model as follows. Firstly, a parent had decided whether to have a first child or not. After the parent observed the first child’s gender, he or she then decided whether to have a second child or not. If he or she had decided not to have a second child, we observe a single-child family. If he or she had decided to have a second child, the mechanism would continue until the parent satisfies with his or her child gender mix.

Let $Y_{mi}^{(s)} = \{0, 1\}$ indicate a binary choice variable representing a decision of having an m^{th} child for parent i . Superscript s represents the decision whether to have a child. Let $Y_{mi}^{(g)}$ be an m^{th} child’s gender for parent $i = 1, 2, \dots, N$. Superscript g represents the child’s gender. If $Y_{mi}^{(s)} = 1$, the parent chose to have a child, and we observe the m^{th} child’s gender, $Y_{mi}^{(g)}$ for all $m = 1, 2, \dots, M$. Figure 1 illustrates child gender mechanism using these notations.

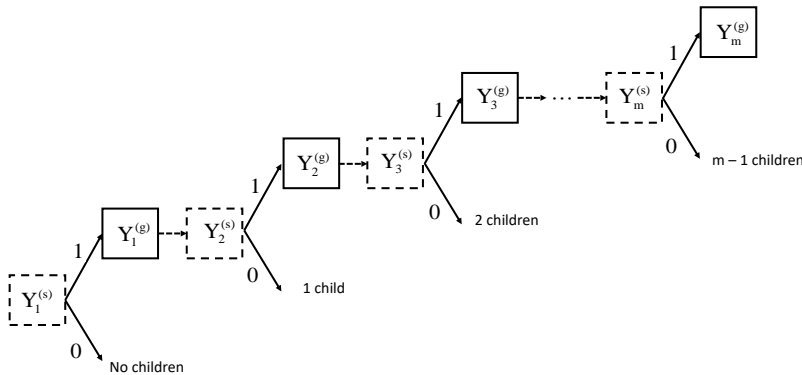


Figure 1: Child decision mechanism

According to the mechanism discussed above, we can build a model to estimate the probability of the parent decision in each stage. Suppose that we observe K characteristics of each parent x_{ik} , where $k = 1, 2, \dots, K$. The probability of observing the parent i had decided to have the m^{th} child, j_m ; the $(m - 1)^{th}$ child, $j_{m-1}; \dots$; the first child, j_1 is

$$\Pr\left(Y_{mi}^{(s)}, Y_{(m-1)i}^{(s)}, \dots, Y_{1i}^{(s)} | Y_{(m-1)i}^{(g)}, Y_{(m-2)i}^{(g)}, \dots, Y_{1i}^{(g)}\right) = p_{mij_m} \times p_{(m-1)ij_{m-1}} \times \dots \times p_{1ij_1}.$$

In this study, we adopted the [5]’s Generalize Maximum Entropy (GME) approach for a discrete choice model to estimate a series of decisions to have children, p_{mij_m} . In the GME framework, the observed noisy data $Y_{mi}^{(s)} = y_{mij_m}$ can be modeled by decomposing into the signal component (unknown probabilities) p_{mij_m} and noise component e_{mij_m} as follows:

$$y_{mij_m} = p_{mij_m} + e_{mij_m} = G(x_i' \beta_{j_m}) + e_{mij_m}, \tag{3.1}$$

where $G(\cdot)$ is a link function linking the probability p_{mij_m} with the linear structure $x_i' \beta_{j_m}$. We have to recover the unknown and unobservable, p_{mij_m} and e_{mij_m} , from the observed noisy data y_{mij_m} and the known covariate x_{ik} . This information of the observed data and covariates can be written as an ill-posed inverse problem with noise as follows:

$$\sum_i x_{ik} y_{mij_m} = \sum_i x_{ik} p_{mij_m} + \sum_i x_{ik} e_{mij_m}, \tag{3.2}$$

for $j_m = 0, 1$ and $k = 1, 2, \dots, K$. The problem is ill-posed because there are $2K$ moment relations and $2N (> 2K)$ unknown parameters.

Conventionally, this ill-posed problem can be solved by imposing parametric restrictions on the link functions that link the unknown probabilities and the covariates [14] or using Bayesian procedures [15]. Under these alternatives, they require a correct statistical model and complete data for valid estimation and testing. The GME approach avoids some of these potential problems of conventional methods.

Using Shannon’s (1948) entropy, measure

$$H(p) = - \sum_j p_j \log p_j, \tag{3.3}$$

[16] proposed maximizing 3.3 subject to appropriate information-moment constraints and normalization probability constraints to recover the unknown probabilities. This is called maximum entropy (ME) principle. For the multinomial choice problem, [17] showed how to recover the multinomial probabilities using the ME principle. [5] showed that maximum likelihood estimation (MLE) of logit model [14] and ME solution for multinomial choice problem are equivalent. The correspondence between ML logit and ME solutions is that the information-moment data constraints in ME formulation are the same as the ML logit first

order conditions. Thus, the ME solution resulting from the entropy maximization subject to the information-moment data constraints has the same mathematical form as logistic distribution.

[5] proposed a generalized maximum entropy formulation for multinomial choice problem by relaxing the strong assumption of the information-moment relations in ME formulation. To solve the inverse-moment relation in (3.2), [5] reparameterized the unobservable e_{mij_m} to be in the form of probability in order to fit in the entropy formalism. Each unobserved component is supposed to be the expected value of a discrete random variable with support $\mathbf{v}_{mij_m} = (v_{mij_m1}, v_{mij_m2}, \dots, v_{mij_mH})'$ and corresponding unknown weights (probabilities) $\mathbf{w}_{mij_m} = (w_{mij_m1}, w_{mij_m2}, \dots, w_{mij_mH})'$, for $H \geq 2$. If we imposed the normalization constraints on the weights, where $\sum_h w_{mij_mh} = 1$, then $e_{mij_m} = \sum_h v_{mij_mh} w_{mij_mh}$ have the properties of probability. Thus equation (3.2) can be rewritten as

$$\sum_i x_{ik} y_{mij_m} = \sum_i x_{ik} p_{mij_m} + \sum_i \sum_h x_{ik} v_{mij_mh} w_{mij_mh}, \quad (3.4)$$

From the principle of ME, p_{ij} that best represents the data must maximize the entropy function

$$\max_{p,w} H(p_{mij_m}, w_{mij_mh}) = - \sum_i \sum_{j_m} p_{mij_m} \log(p_{mij_m}) - \sum_i \sum_{j_m} \sum_h w_{mij_mh} \log(w_{mij_mh})$$

subject to constraints (3.4) and the following normalization constraints

$$\begin{aligned} \sum_{j_m} p_{mij_m} &= 1, \quad \forall i = 1, \dots, N \\ \sum_h w_{mij_mh} &= 1, \quad \forall i = 1, \dots, N, \forall j_m = 0, 1. \end{aligned}$$

We can solve this problem by setting up the corresponding Lagrangian as follows:

$$\begin{aligned} L = & - \sum_i \sum_{j_m} p_{mij_m} \log(p_{mij_m}) - \sum_i \sum_{j_m} \sum_h w_{mij_mh} \log(w_{mij_mh}) \\ & + \lambda_{10} \left(\sum_i x_{i1} p_{mi0} + \sum_i \sum_h x_{i1} v_{mi0h} w_{mi0h} \right) + \dots \\ & + \lambda_{k0} \left(\sum_i x_{ik} p_{mi0} + \sum_i \sum_h x_{ik} v_{mi0h} w_{mi0h} \right) \\ & + \lambda_{11} \left(\sum_i x_{i1} p_{mi1} + \sum_i \sum_h x_{i1} v_{mi1h} w_{mi1h} \right) + \dots \\ & + \lambda_{k1} \left(\sum_i x_{ik} p_{mi1} + \sum_i \sum_h x_{ik} v_{mi1h} w_{mi1h} \right) \end{aligned}$$

$$\begin{aligned}
 & + \rho_1 \left(1 - \sum_{j_m} p_{m1j_m} \right) + \dots + \rho_N \left(1 - \sum_{j_m} p_{mNj_m} \right) \\
 & + \delta_{10} \left(1 - \sum_h w_{m10h} \right) + \dots + \delta_{N0} \left(1 - \sum_h w_{mN0h} \right) \\
 & + \delta_{11} \left(1 - \sum_h w_{m11h} \right) + \dots + \delta_{N1} \left(1 - \sum_h w_{mN1h} \right).
 \end{aligned}$$

The Lagrangian’s first-order conditions form a basis for solving the optimization problem and for recovering the unknown probability \hat{p}_{mij_m} , \hat{e}_{mij_m} , and $\hat{\lambda}_{kj_m}$. Notice that parameter $\hat{\beta}_{kj_m}$ equate $-\hat{\lambda}_{kj_m}$. Solving the first-order conditions give the solution

$$\hat{p}_{mij_m} = \frac{\exp \left(\sum_k x_{ik} \hat{\beta}_{j_m} \right)}{\sum_{j_m} \exp \left(\sum_k x_{ik} \hat{\beta}_{j_m} \right)},$$

and

$$\hat{w}_{mij_mh} = \frac{\exp \left(\sum_k \sum_h x_{ik} \hat{\beta}_{j_m} v_{mij_mh} \right)}{\sum_{j_m} \exp \left(\sum_k \sum_h x_{ik} \hat{\beta}_{j_m} v_{mij_mh} \right)}.$$

4 Data

The data used in this study are from the 2015 Household Socio-Economic Survey (SES) collected by Thailand’s National Statistical Office. The survey collected data on 125,336 individuals in 43,400 households in Thailand. To examine child-gender preference, which relies on the probability of parents having an additional child given the gender of the previous child, this study matched parents with their children in each of the household. This study examined child bearing decision and unmarried individuals may behave differently. Moreover, studying the fathers’ and the mothers’ preferences would be redundant. This study, thus, only examines married mothers’ decision to have additional children.

As the SES collected data of only individuals living inside the household, the data do not to capture children living outside the household. Since the choice to leave the household can be endogenous, this study limited the age of mother to 15-30 years old. The study also eliminated samples that mothers had children before the age of 15 years old. As a result, the oldest child in the sample could not be more than 15 years old and is unlikely to leave the household. After the restrictions, the final sample size is 3,826 mothers. All variable descriptions and sample statistics are provided in Table 1.

Table 1: Data Description

Variable	Description	N	Mean	SD
ch_d1	Dummy (=1 if have first child)	3826	0.649	0.477
ch_d2	Dummy (=1 if have second child)	3826	0.244	0.430
ch_d3	Dummy (=1 if have third child)	3826	0.037	0.188
ch_d4	Dummy (=1 if have forth child)	3826	0.005	0.074
xage	Age of the mother	3826	25.508	3.720
xedu_hs	Dummy (=1 if have high school degree)	3826	0.504	0.500
xedu_c	Dummy (=1 if have collage degree)	3826	0.139	0.346
xurban	Dummy (=1 if live in urban area)	3826	0.578	0.494
ch_G1	Dummy (=1 if first child is girl)	2477	0.512	0.500
ch_G2	Dummy (=1 if first two children are both girls)	913	0.239	0.427
ch_G3	Dummy (=1 if first three children are all girls)	131	0.122	0.329
ch_GB2	Dummy (=1 if first two children are mixed genders)	913	0.493	0.500
ch_GB3	Dummy (=1 if first three children are mixed genders)	131	0.718	0.452

5 Results

This study measures child-gender preference by examining the probability of having an additional child given genders of the previous children using the logit and entropy based models. This requires estimating a system of regression equations representing the decision to have the first, second, third and forth child. Specifically, equations (L1) and (E1) estimate the mother's decision to have the first child ch_d1 given her characteristics $x = \{xage, xedu_hs, xedu_c, xurban\}$ using the logit and entropy based models, respectively. Equations (Lj) and (Ej) estimate the mother's decision to have the j^{th} child ch_dj given her characteristics x and gender of her previous children ch_G and ch_GB , where ch_G is the dummy for all previous children being girls and ch_GB is the dummy for previous children having mixed genders. To examine the issue of child-gender preference in more details, interaction terms were added in the models. That is, i_hs_G is the interaction term between $xedu_hs$ and ch_G and i_hs_GB is the interaction term between $xedu_hs$ and ch_GB . In particular, the i_hs_G allows us to separately study the existence of girl preference among mothers with and without high school degree. The i_hs_GB allows us to separately study the existence of mixed gender preference among mothers with and without high school degree.

The estimations for the logit model (L1)-(L4) and the entropy model (E1)-(E4) are shown in Table 2. It should be noticed that, for the (L1) and (E1) estimation, the number of observation is not small. Without the interaction term between previous child's gender and mother's education, the estimation has no multicollinearity problem. Therefore, the logit and entropy estimates are not significantly different. However, for the estimations of (L4) and (E4), the number of observation is reduced to 131 individuals as there were not many mothers having three children or more. In addition, the equations include two interaction terms between previous children's gender and the mother's educational level causing

Table 2: The Logit and GME estimation

	(L1)	(L2)	(L3)	(L4)	(E1)	(E2)	(E3)	(E4)
xage	0.09*** (0.01)	0.21*** (0.01)	0.09** (0.04)	0.15 (0.13)	0.09*** (0.01)	0.21*** (0.01)	0.09** (0.04)	0.13 (0.12)
xedu_hs	0.05 (0.08)	-0.38*** (0.13)	-0.74** (0.36)	16.15 (1385.77)	0.05 (0.08)	-0.38*** (0.13)	-0.73** (0.36)	3.01 (1.92)
xedu_c	-0.83*** (0.11)	-0.67*** (0.16)	-0.76* (0.46)	0.79 (1.04)	-0.82*** (0.11)	-0.67*** (0.16)	-0.75 (0.45)	0.76 (1.02)
xurban	-0.50*** (0.07)	-0.11 (0.09)	-0.27 (0.19)	-0.46 (0.52)	-0.50*** (0.07)	-0.11 (0.09)	-0.26 (0.19)	-0.42 (0.50)
ch_G		-0.02 (0.12)	-0.72** (0.33)	14.84 (1385.77)		-0.02 (0.12)	-0.71** (0.33)	1.79 (2.00)
i_hs_G		-0.17 (0.18)	0.98* (0.53)	-16.26 (1385.77)		-0.17 (0.18)	0.97* (0.53)	-3.12 (2.44)
ch_GB			-0.65** (0.27)	15.25 (1385.77)			-0.64** (0.27)	2.16 (1.75)
i_hs_GB			0.69 (0.45)	-15.81 (1385.77)			0.68 (0.45)	-2.70 (2.01)
constant	-1.37*** (0.24)	-5.59*** (0.38)	-3.39*** (1.01)	-21.00 (1385.77)	-1.37*** (0.24)	-5.58*** (0.38)	-3.35*** (1.00)	-7.35 (3.71)
N	3826	2477	913	131	3826	2477	913	131

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

multicollinearity problem. As a result, the entropy estimates have much lower standard errors as the maximum entropy is by nature robust to multicollinearity and small sample-size problems. Comparing the estimation of the entire systems (L1)-(L4) and (E1)-(E4), the entropy based method outperformed the logit based method. It also should be noted that the logit and entropy based models show similar estimation outcomes, which means that the results are robust to the different estimation methodologies.

As the entropy based model provides lower standard deviations than the Logit model, especially when multicollinearity problem is presented, the equations (E1m.30)-(E4m.30) in Table 3 shows the marginal effects of each regressor on the mother's probability of having an additional child. As the SES collected data of only individuals living inside the household and this study limited the age of mother to 15-30 years old to prevent the bias from children's endogenous decision to leave home. The restriction on the mothers' age may raise a question as a 30 years old mother may still have more children later on. Therefore, for the robustness check, this study also examines the decisions to have an additional child of mothers from 15-45 years old and the marginal effects are shown in (E1m.45)-(E4m.45) in Table 3.

From the (E2m.30) and (E4m.30), gender of the first child has no significant effect on the mother's decision to have the second child and genders of the previous children have no effect on the mother's decision to have the forth child. However, from the (E3m.30), ch_GB is negatively correlated with the decision to have the third child. Moreover, the coefficient for ch_G is significantly negative and the coefficient for i_{hs_G} is significantly positive.

This indicates that the decision to have the second and forth child does not depend on gender of previous children. However, the decision to have the third child depends significantly on the gender mix of the first two child. Specifically, the results show an evidence for mixed-gender preference. Once the parents have a boy and a girl, they are more likely to stop having children. In addition, mothers without high school degree tend to have less children if their first child is a girl and mothers with high school degree tend to have more children of their first child is a boy. This should imply a girl preference among mothers without high school degree and a boy preference among mothers with high school degree.

In addition to the gender of previous children, the decision to have an additional child depends on mothers' age, education and living location. Since the age restriction is low, mothers with higher age tend to have more children. Mothers with higher education and live in urban area tend to have few children.

As for the robustness check over the age restrictions of mothers, the estimations for the two age ranges of mothers provide similar results for child-gender preference with an exception of the effect of the previous two children being female on the decision to have the third child. With the 15-30 years old age restriction, the results show a significant girl preference. However, with the 15-45 years old age restriction, the results show an insignificant effect. It should be noted that, with the data limitation, the estimation faces a risk of biases in all mothers' age ranges. When using the data of younger mothers, the mothers can potentially have more

Table 3: The marginal effects of the GME estimation for the decisions to have an additional child for (1) mothers aged 15-30 years old and (2) mothers aged 15-45 years old

	(E1m_30)	(E2m_30)	(E3m_30)	(E4m_30)	(E1m_45)	(E2m_45)	(E3m_45)	(E4m_45)
xage	0.02*** (0.00)	0.04*** (0.00)	0.01** (0.00)	0.17 (0.02)	0.01*** (0.00)	0.01*** (0.00)	0.00*** (0.00)	0.00* (0.00)
xedu_hs	0.01 (0.02)	-0.08*** (0.03)	-0.09** (0.04)	0.44** (0.21)	0.01*** (0.01)	-0.01 (0.02)	-0.07*** (0.02)	-0.00 (0.07)
xedu_c	-0.20*** (0.03)	-0.13*** (0.03)	-0.08** (0.04)	0.12 (0.18)	-0.06*** (0.01)	-0.03** (0.02)	-0.06*** (0.02)	0.01 (0.05)
xurban	-0.11*** (0.02)	-0.02 (0.02)	-0.03 (0.02)	-0.05 (0.06)	-0.08*** (0.01)	-0.02** (0.01)	-0.02* (0.01)	-0.06** (0.03)
hch_G		-0.00 (0.03)	-0.09** (0.04)	0.23 (0.26)	0.00 (0.01)	0.00 (0.02)	-0.01 (0.02)	0.01 (0.06)
i_hs_G		-0.03 (0.04)	0.12* (0.07)	-0.40 (0.31)	-0.03 (0.02)	-0.03 (0.02)	0.06* (0.03)	-0.15 (0.12)
hch_GB			-0.08** (0.03)	0.28 (0.23)			-0.05*** (0.02)	-0.01 (0.04)
i_hs_GB			0.09 (0.06)	-0.35 (0.26)			0.04 (0.03)	-0.02 (0.08)
constant	-0.30*** (0.05)	-1.17*** (0.07)	-0.43*** (0.13)	-0.94** (0.47)	-0.05** (0.02)	-0.27*** (0.03)	-0.27*** (0.04)	-0.33*** (0.10)
N	3826	2477	913	131	14337	10342	5073	945

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

children later on in their lives. When using the data of older mothers, their children are also older. The children can selectively leave the household and we cannot observe children not living in the same household in this data set. The misspecification due to mothers having children later on should not be correlated with gender of their previous children. However, the misspecification due to the children's decision to leave home can be correlated with their gender. Therefore, to examine the child-gender preference, this study focuses on the results from the estimation using the data of younger mothers.

6 Conclusions

The results from this study present different types of child-gender preferences among Thai population. Mothers with lower education prefer girls, while mothers with higher education prefer boys. However, mothers with all levels of education prefer mixed genders. That is, Thai family prefer to have at least one boy and one girl.

The results from this study are different from those from China, India and Vietnam that show evidence of boy preference [8]. However, the results are consistent with [6], which states that the boy preference culture diminishes overtime in South Korea. For the case of Thailand, we only have an evidence for a mild boy preference [11] and some mixed-gender preference [13, 12, 9]. The results of this study confirm the results of previous studies that Thai parents, in general, have a mixed-gender preference and no strong evidence of boy or girl preference in the overall population.

It should be noted that, with the data limitation, this study did not control for potential different costs of raising a boy and a girl. As the decision to have an additional child can be sensitive to the cost difference, the estimation of child-gender preference using the decision to have an additional child can potentially face this bias, especially among mothers with lower education and income. In addition, this study estimated a system of child bearing decisions separately. An algorithm to simultaneously estimated the entire decision system would yield a more efficient estimation.

Acknowledgements : We acknowledge the partial support of the Center of Excellence in Econometrics, Faculty of Economics, Chiang Mai University, Thailand. We also would like to thank Prof. Colin McKenzie from Keio University for contributing valuable ideas and comments for this study.

References

- [1] R. Takaku, First Daughter Effects in Japan (2016) Available at Social Science Research Network: <https://ssrn.com/abstract=2883069>.

- [2] L. Li, X. Wu, Gender of children, bargaining power, and intrahousehold resource allocation in China, *The Journal of Human Resources* 46 (2) (2011) 295-316.
- [3] A. Ichino, E. A. Lindstrom, E. Viviano, Hidden consequences of a first-born boy for mothers, *Economics Letters* 123 (3) (2014) 274-278.
- [4] T. M. Vu, The lexicographic preference for a son: evidence from household data in Vietnam (2012) Available at Osaka School of International Public Policy Discussion Paper: <http://www.osipp.osaka-u.ac.jp/archives/DP/2012/DP2012E001.pdf>.
- [5] A. Golan, G. Judge, J. M. Perloff, A maximum Entropy approach to recovering information from multinomial response data, *Journal of the American Statistical Association* 91 (434) (1996) 841-853.
- [6] S. H. Yoo, S. R. Hayford, V. Agadjanian, Old habits die hard? lingering son preference in an era of normalizing sex ratios at birth in South Korea, *Population Research and Policy Review* 36 (1) (2017) 25-54.
- [7] J. Bongaarts, The implementation of preferences for male offspring, *Population and Development Review* 39 (2) (2013) 185-208.
- [8] J. Bongaarts, C. Z. Guilmoto, How many more missing women? excess female mortality and prenatal sex selection, 1970-2050, *Population and Development Review* 41 (2) (2015) 241-269.
- [9] V. Prachuabmoh, J. Knodel, J. O. Alers, Preference for sons, desire for additional children, and family planning in Thailand, *Journal of Marriage and Family* 36 (3) (1974) 601-614.
- [10] J. Knodel, P. Pitaktesombati, Fertility and family planning in Thailand: results from two rounds of a national study, *Studies in Family Planning* 6 (11) (1975) 402-413.
- [11] J. Knodel, V. Prachuabmoh, Preferences for sex of children in Thailand: a comparison of husbands' and wives' attitudes, *Studies in Family Planning* 7 (5) (1976) 137-143.
- [12] K. Wongboonsin, V. P. Ruffolo, Sex preference for children in Thailand and some other South-East Asian countries, *Asia-Pacific Population Journal* 10 (3) (1995) 43-62.
- [13] J. Knodel, V. P. Ruffolo, P. Ratanalangarn, K. Wongboonsin, Reproductive preferences and fertility trends in post-transition Thailand, *Studies in Family Planning* 27 (6) (1996) 307-318.
- [14] D. L. McFadden, Conditional logit analysis of qualitative choice behavior, *Frontiers in Econometrics*, Academy Press, New York (1974), 105-142.
- [15] J. H. Albert, S. Chib, Bayesian analysis of binary and polychotomous response data, *Journal of the American Statistical Association* 88 (1993) 669-679.

- [16] E. T. Jaynes, Information theory and statistical mechanics, *Physical Review* 106 (4) (1957) 620-630.
- [17] E. S. Soofi, A generalizable formulation of conditional logit with diagnostics, *Journal of the American Statistical Association* 87 (419) (1992) 812-816.

(Received 31 August 2017)

(Accepted 30 October 2017)