Educated Iranian Women in Favor of Having Girls: CART Classification Approach

Mahsa Saadati[1] ,Arezoo Bagheri[2]

[1,2] National Population Studies & Comprehensive Management Institute, Tehran, Iran
[1] Email: mahsa.saadati@gmail.com, [2] Corresponding author's Email: abagheri_000@yahoo.com

Abstract

One of the three principal components of population dynamics, which determine the size, structure, and composition of the population, is fertility. Some determinants such as sex preferences for children may have substantial implications in a couple's fertility behavior. Due to consequences of sex preferences, it has been a great concern to many researches and policy makers. The main aim of this article is to investigate whether parents have preferred the gender of their children, and if so what are the most influential factors on this variable. In this article, data from 1250 ever married Iranian women age 15-49 are used to classify women's sex preferences for children. An efficient decision tree algorithm as Classification & Regression Trees for classification of binary response variables is applied to study whether parents prefer one sex over the other. Age, educational level, place of residence, and difference number of siblings for women, were nominated as predictors while only women's age, educational level and difference number of siblings were appeared in extracted decision tree. The results indicated that educated Iranian women in both young and old ages preferred to have girls instead of boys.

***Keywords*:** Sex Preferences for Children, Decision Tree, Classification & Regression Trees Algorithm.

1.Introduction

Globally, sex preferences for children have been a prominent issue in demographic study because of its potential negative social and demographic implications (Fuse, 2010). Son preference is prevalent in many parts of the world particularly South and East Asia, parts of the Middle East, and North Africa. In Bangladesh and Nepal, more than 95 percent of parents preferred to have a son. Similarly, in Burkina Faso and Senegal, more than 30 percent of women had a son preference. Son preference has been documented in India through sex-selective abortions that lead to abnormally high sex ratios at birth (Arnold, 1992), contraceptive use behavior and gender bias in the allocation of food and health care (Arokiasamy, 2002).

Daughter preference has also been observed in some West African countries including Ghana (21.3 percent), Malawi (21.2 percent) and Liberia (22.2 percent) (Fuse, 2010).

Arnold (1997) provides a detailed study of Demographic and Health Surveys (DHS) for 44 countries with in the period from 1986 to 1995. He found "son preference in a range of different countries, demonstrating that such preferences do not emerge from a single set of historical and cultural experiences". While the Southeast Asian nations do not show any consistent gender preference, the Caribbean is the only region studied by Arnold (1999), where a prevalent preference for daughters has been found. In general, however, he argued that the effect of gender preferences on fertility and family planning is not very strong.

In Asia, the preference of many parents for sons over daughters has led to some 80 million girls missing from what should be the normal balance between men and women in a society, perhaps because they have been aborted, neglected, or directly killed (Francis, 2015).

The arrival of prenatal sex determination in the early 1980s combined with age old preference for sons have resulted in unnaturally high sex ratios (males to females) in a number of Asian countries, including the world's two most populous, China and India (Scharping, 2003; Zhu et al., 2009; Chung and Das Gupta , 2007).

Kugler and Kumar (2015) studied whether Indian parents make trade-offs between the number of children by using data from nationally representative household surveys. Given a strong son preference in India, parents tend to have more children if the first born is a girl. The preference for male children in India means that when a household has a first born who is a girl, parents will continue to have more children until they have a boy. Thus, in a son biased society, the first child's sex should be a good predictor of the probability of having a second child or of the total number of children in the household. Having a first-born who is a female is strongly positively correlated with family size.

Edlund and Lee (2013) stated that there is a very slight tendency for mothers in good condition to have more boys on South Korea which sex ratios increased significantly with the mother's educational level.

American parents, especially fathers, also favor boys over girls. This preference for sons is less severe and subtler than in Asia, but it has consequences nonetheless. Dahl and Moretti (2008) showed how the American parental preference affects divorce, child custody, marriage, and shotgun marriage when the sex of the child is known before birth. They found that the bias for boys is quantitatively important. The authors concluded that the preference for boys of American parents seems to be largely driven by fathers. Though, women had only a slight preference for daughters. Also they noted that this preference for boys could matter

more in the future. More importantly, the bias for boys evidenced by their results may lead to worse outcomes for daughters.

McDougall et al. (1999) examined parental preferences for sex of children in Canada using data from the 1984 Canadian Fertility Survey. Attitudinal and birth timing measures were used to assess sex preferences at different parities. Both measures indicated that the primary preference is for at least one child of each sex. The birth timing measure indicated that neither boys nor girls are preferred as first-born children by women and their husbands. The attitudinal measure showed that sons are preferred as first-born children among women with a sex preference. However, the greater percentage of zero-parity women has no sex preference for their first-born child. Furthermore, the greater percentage of women at every parity expresses no sex preference.

Andersson et al. (2007) compared childbearing dynamics and possible underlying sex preferences among Finnish-born immigrants in Sweden and members of the Swedish-speaking minority in Finland. For Finland, they observed a continuous boy preference among the national majority and the Swedish-speaking minority as reflected in higher third-birth rates of mothers of two girls than of mothers of two boys. Evidence of similar preferences was found for Finnishborn migrants in Sweden, where the native-born population instead appeared to have developed a girl preference.

Seidl (1995) stated that development of sex preferences for children has also been in some studies on advanced western societies (Diekmann and Schmidheiny, 2004; Lundberg, 2005; Raley and Bianchi, 2006). This development might, first, be due to medical advances that facilitate parents' deliberate choice of their child's sex (Dahl et al. 2006). Secondly, it has been argued that in modern low-fertility societies, the influence of the sex composition of previous children on couples' childbearing behavior might intensify, because factors affecting the decision to have another child should become more important (Gray and Evans 2005).

Although it has been claimed that changes in a society's gender system may lead to a decreasing effect of children's sex on parents' fertility decisions (Pollard and Morgan 2002), recent studies have shown that modernization and increasing gender equality do not necessarily bring about parental gender indifference. Andersson et al. (2006) showed that, on the one hand, new sex preferences (in favor of girls) are likely to have evolved in Denmark, Norway, and Sweden in the 1980s, while, on the other hand, culturally rooted preferences for sons in Finland have been left unaffected by improvements in women's social, political, and reproductive rights.

Hank and Kohler (2000) analyzed gender preferences for children from a cross-national perspective of 17 European countries. Their results basically supported the findings of older studies dealing with gender preferences in Western societies. There was a strong tendency towards a preference for a mixed sex composition. However, they found some unexpected indication for a girl preference in the Czech Republic, Lithuania, and Portugal.

Strong gender preferences, combined with infanticide, sex-selective abortions, or sex selection technologies, may lead to a serious distortion in the natural sex ratio (Park and Cho, 1995). Such an imbalance between the two sexes could, for example, cause a delay in the age of marriage, or an increase in the number of people who never marry. Furthermore, gender preferences may have substantial implications for a couple's fertility behavior. One might assume that parents who desire one or more children of a certain sex may have larger families than the others. Parents who fail to achieve the desired sex balance (or ratio) by the time they reach the number of children intended, might tend to revise their family size goals upward (Gray and Morrison, 1974).

Sex preference can also lead to gender bias in the allocation of food and health care (Mishra et al., 2004). Birth intervals have been observed to be largest for women having equal number of boys and girls, intermediate for those having more boys than girls and shortest for those having more girls than boys. Campbell and Campbell (1997) reported that the preference for sons contributed to the high fertility levels in the western area of Sierra Leone. This may imply that sex preferences tend to increase fertility levels.

Some Iranian researchers like Shahbaziyan et al. (2014) investigated that whether sex preferences of parents, more specifically mothers, can change the family fertility behavior in Kangavar country, Iran. They studied fertility according to sex preferences, educational level, job status and age of women. They found that there are a high relationship between educational level and gender preference for children. They showed that more sons led to less fertility. Mansurian and Khushnevis (2006) studied the sex preferences influence of ever married women on their fertility behavior in Tehran, Iran. They found that fertility behavior in families with more boys is lower than families with more girls. The women in this study had preference to have more boys. Hejazi (2013) studied attitudes of employed women in Isfehan province to have second child. They have considered women's age, educational level, sex preferences, and job status as predictors. They didn't find any sex preferences in their study.

One of the important determinants of fertility behaviors is sex preferences for children and up to our knowledge, there aren't any studies which consider Iranian women's sex preferences

for children as a response variable and determine directly its influential factors. Therefore, the main aim of this article is to investigate factors which affect on Iranian women's sex preferences for children. To do so, this article arranged as follows; Section (2), materials and methods, introduces an applicable classification tree method as Classification & Regression Trees algorithm (CART) and describes data. The results of applying the defined classification algorithm on Iranian case study presented in Section (3). Section (4) indicates the conclusion of this application.

2. Material and Methods

Most of researchers modeled sex preferences by logistic regression. There are a number of reasons for difficulties of traditional statistical methods like logistic regression to investigate sex preferences. First, there are generally many possible predictor variables which make the task of variable selection difficult. Logistic regressions are poorly suited for this sort of multiple comparisons. Second, complex interactions or patterns may exist in the data. For example, the value of one variable may substantially affect the importance of other variables. These types of interactions are generally difficult to model, and virtually impossible to model when the number of interactions and variables becomes substantial. Third, a common, but incorrect, method of handling missing data in most of traditional statistical methods is to exclude cases with missing values; this is both inefficient and runs the risk of introducing bias in the analysis.

To model sex preferences, classification and regression trees that are machine-learning methods for constructing prediction models from data could be used. The models are obtained by recursively partitioning the data space and fitting a simple prediction model within each partition. As a result, the partitioning can be represented graphically as a decision tree. It is the target variable that determines the type of decision tree needed. If the target variable is categorical or contiguous, it is considered as classification or regression trees.

A decision problem includes following components (Timofeev, 2004):

- A categorical outcome or dependent variable;
- Independent variables;
- Learning dataset;
- Test or future dataset;
- A prior probability for each outcome, which represents the probability that a randomly-selected future case will have a particular outcome.

- A decision loss or cost matrix that represents the inherent cost associated with misclassifying a future case.

The theory of a decision tree has the following main parts: a root node is the starting point of the tree; branches connect nodes showing the flow from question to answer. Nodes that have child nodes are called interior nodes. Leaf or terminal nodes are nodes that do not have child nodes and represent a possible value of target variable given the variables represented by the path from the root.

The basic idea of decision tree analysis is to split the given source data set into subsets by recursive portioning of the parent node into child nodes based on the homogeneity of within-node instances or separation of between-node instances with respect to target variables. For each node, attributes are examined and the splitter is chosen to be the attribute such that after dividing the nodes into two child nodes according to the value of the attribute variable, the target variable is differentiated to the best using algorithm. Because of this, it is needed to be able to distinguish between important attributes, and attributes which contribute little to overall decision process. This process is repeated on each child node in a recursive manner until splitting is either non-feasible or all certain pre-specified stopping rules are satisfied.

2.1. CART Algorithm

Several statistical algorithms for building decision trees are available, including CART (Classification and Regression Trees) (Breiman et al., 1984), C4.5 (Quinlan , 1993), CHAID (Chi-Squared Automatic Interaction Detection) (Kass, 1980) and QUEST (Quick, Unbiased, Efficient, Statistical Tree) (Loh and Shih, 1997). Table (1) provides a brief comparison of the four most widely used decision tree methods (Bhukya and Ramachandram, 2010; Lin et al., 2006).

Table 1. Comparison of Different Decision Tree Algorithms

Methods	CART	C4.5	CHAID	QUEST
Measure used to select input variable	Gini index; Towing criteria	Entropy info-gain	Chi-square	Chi-square for categorical variables; j-way ANOVA for continuous/ordinal variables
Pruning	Pre-pruning using a single pass algorithm	Pre-pruning using a single pass algorithm	Pre-pruning using chi-square test for independence	Post-pruning
Dependent variable	Categorical/ Continuous	Categorical/ Continuous	Categorical	Categorical
Input variables	Categorical/ Continuous	Categorical/ Continuous	Categorical/ Continuous	Categorical/ Continuous
Split at each node	Binary; Split on linear combinations	Multiple	Multiple	Binary; Split on linear combinations

CART analysis consists of four following basic steps:

- The first step consists of tree building, during which a tree is built using recursive splitting of nodes. Each resulting node is assigned a predicted class, based on the distribution of classes in the learning dataset which would occur in that node and the decision cost matrix. The assignment of a predicted class to each node occurs whether or not that node is subsequently split into child nodes.
- The second step consists of stopping the tree building process. At this point a maximal tree has been produced, which probably greatly overfits the information contained within the learning dataset.
- The third step consists of tree pruning, which results in the creation of a sequence of simpler and simpler trees, through the cutting off on increasingly important nodes.
- The fourth step consists of optimal tree selection, during which the tree that fits the information in the learning dataset, but does not overfit the information, is selected from the sequence of pruned trees.

Complexity and robustness are competing characteristics of models that need to be simultaneously considered whenever building a statistical model. The more complex a model is, the less reliable it will be when it is used to predict future records. An extreme situation is to build a very complex decision tree model that spreads wide enough to make the records in each leaf node 100 percent pure that means all records have the target outcome. Such a decision tree would be overly fitted to the existing observations and have few records in each leaf, so it could not reliably predict future cases and, thus, would have poor generalizability which means lack robustness.

Cross validation is a computationally-intensive method for validating a procedure for model building, which avoids the requirement for a new or independent validation dataset. In cross validation, the learning dataset is randomly split into *N* sections, stratified by the outcome variable of interest. This assures that a similar distribution of outcomes is present in each of the *N* subsets of data. One of these subsets of data is reserved for use as an independent test dataset, while the other *N*-1 subsets are combined for use as the learning dataset in the model-building procedure. The entire model-building procedure is repeated *N* times, with a different subset of the data reserved for use as the test dataset each time. Thus, *N* different models are produced, each one of which can be tested against an independent subset of the data. Cross validation is based is on the amazing fact that the average performance of these *N* models is an excellent estimate of the performance of the original model on a future independent set of cases.

Although the CART algorithm manual recommends experimenting with different splitting criteria, these criteria will give identical results if the outcome variable is a binary categorical variable. The most common splitting function is the Gini, followed by Twoing (Timofeev, 2004). Gini index is an impurity-based criterion that measures the divergences between the probability distributions of the target variable's values. The Gini index has been used in various works (Breiman et al., 1984; Gelfand et al., 1991) which is defined as:

$$Gini(y,S) = 1 - \sum_{c_j \epsilon dom(y)} \left(\frac{\left|\sigma_{y=c_j} S\right|}{|S|}\right)^2 \quad (1)$$

where S is a training set and y is the probability vector of the target variable. Consequently, the evaluation criterion for selecting the attribute a_i is defined as:

$$GiniGain(a_i,S) = Gini(y,S) - \sum_{v_{i,j} \epsilon dom(a_i)} \left(\frac{\left|\sigma_{a_i=v_{i,j}} S\right|}{|S|}\right) Gini\left(y, \sigma_{a_i=v_{i,j}} S\right) \quad (2)$$

2.2. Data Description

1250 15-49 year-old ever married women who have had child (children) and intent to have more child (children) in "Childbearing Attitudes and Its Social, Economical and Cultural Factors" survey (Kazemipour, 2014) were considered as data in this study. Sex preferences of these women were indicated by two different questions which measured the number of their CEB and desired number of children. Multistage stratified sampling in this study used to select women who referred in public health and treatment centers to vaccinate their children in 31 provinces in Iran, 2014. Samples have been chosen randomly by Probability Proportional Sampling to Size (PPS) sampling method.

Different factors may affect women's sex preference. The list of dependent (response) and independent variables (factors) used in this study are as follows:

Sex preferences for children (response variable): From 6231 women age 15-49 in the survey (Kazemipour, 2014); those didn't want any more children and were childless deleted from interested population. Then, the women's number of ever born and desired boys summed up and the same index calculated for girls as well in the resulted population in the previous step. The difference between the sum of girls and boys computed as preference values. Sex preferences were categorized as *boy preference*, *girl preference*, and *no sex preference*. All respondents with zero preference values were classified as having no preference for a child's sex and deleted from interested population. All respondents with negative preference values were classified as having preference for boys and all those women with positive preference values were classified as having preference for girls. Thus

respondents with boy or son preferences were made the final sample in this study which was 1250 women.

Place of residence: It is a place that women were living in the study time that could be even urban or rural areas.

Women's age: A four categorical variable with levels of 10-19, 20-29, 30-39, and 40-49 that is considered to measure the age of women in survey time.

Women's educational level: It was considered as a categorical variable with five categories of under secondary, high school and diploma, associate and bachelor, master and above, and religious degree.

Women's difference number of siblings: It has three categories as equal, more sisters, and more brothers which represented women with equal number of sisters or brothers, more sisters compared to brothers or more brothers compared to sisters, respectively.

3. Results

The main aim of this study is to classify gender preferences of 15-49 year-old Iranian women. Figure (1) presents pie chart of women's sex preference. 51.04 and 48.96 percent of women preferred to have boys and girls, respectively. Figure (2) shows bar chart of predictor variables in this study. 68.6 percent of women are living in the urban area. 54.9 percent of them are 20-29 years old. Women's educational level for 74.6 percent is diploma and less. 43.3 percent of women have more sisters than brothers while 38.7 percent of women are vice versa.

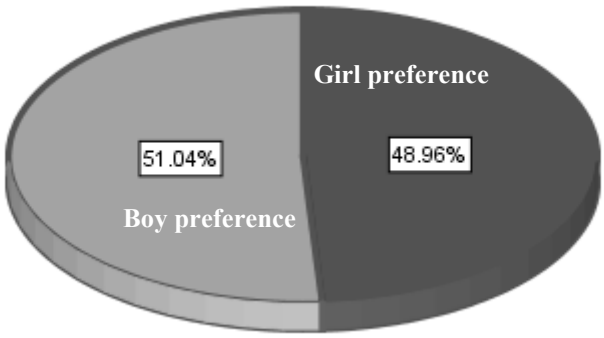

Figure1. Pie Chart of Women's Sex Preference

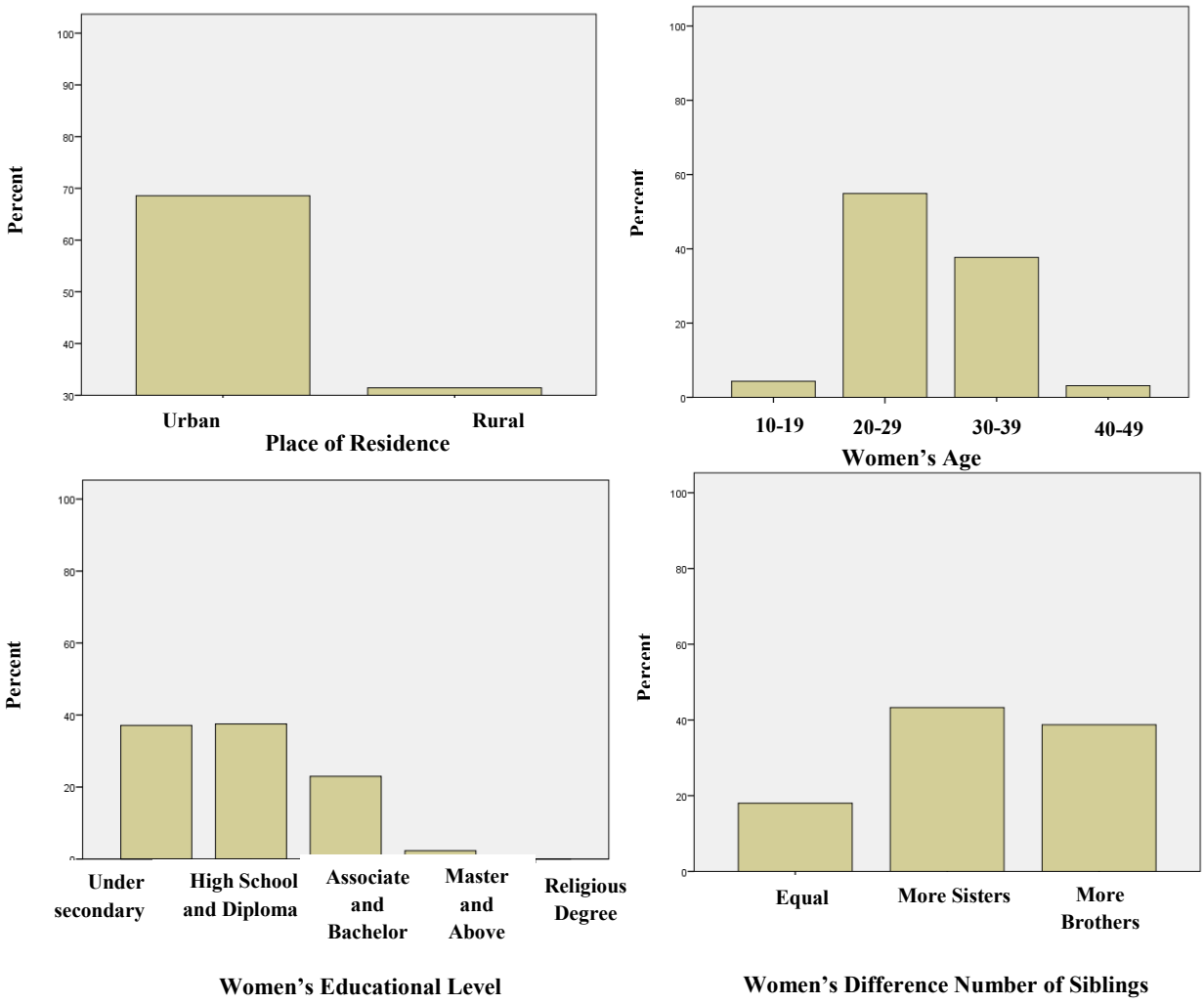

Figure 2. Bar Charts of Independent Variables

Table (2) shows women's sex preferences crossed by predictors in this study. According to the results of this table, women who lived in urban area (50.3 percent), had more than 30 years old (52.7 percent in age 30-39 and 53.8 percent in age 40-49), with educational level of higher than associate and bachelor (51.9 percent of associate and bachelor levels and 58.6 percent of master and above levels), had equal siblings or more sisters (51.1 percent of equal number of siblings and 50.8 percent of more sisters) preferred girls. While the other women in each level of predictors preferred boys.

Table 2. Women's Gender Preference by Predicted Variables

Variables		Children preference (Response Variable)			Chi-Square Test	Asymp. Sig. (2-sided)
Name	Value	Girl	Boy	Total		
Place of Residence	Urban	50.3	49.7	100	1.935	<0.164
	Rural	46.1	53.9	100		
Women's Age	10-19	42.6	57.4	100	5.289	<0.152
	20-29	46.6	53.4	100		
	30-39	52.7	47.3	100		
	40-49	53.8	46.2	100		
Women's Educational Level	Under secondary	45.9	54.1	100	4.876	<0.300
	High school & diploma	49.7	50.3	100		
	Associate and Bachelor	51.9	48.1	100		
	Master & above	58.6	41.4	100		
	Religious Degree	0.0	100.0	100		
Women's Difference Number of Siblings	Equal	51.1	48.9	100	3.027	<0.220
	More sisters	50.8	49.2	100		
	More brothers	45.9	54.1	100		

Figure (3) presents decision trees of women's sex preference according to the selected predictors in this study. The rules that can be extracted from this tree are as follows:

- Women whose age were 10 to 30 years old and their educational levels were associate and above or diploma and lower including religious degree preferred to have girls or boys, respectively.
- Women whose age were 30 to 49 years old and they had equal number of siblings and more sisters preferred to have girls.
- Women whose age were 30 to 49 years old and had more brothers according to their educational levels that were under secondary or high school and above preferred to have boys or girls, respectively.

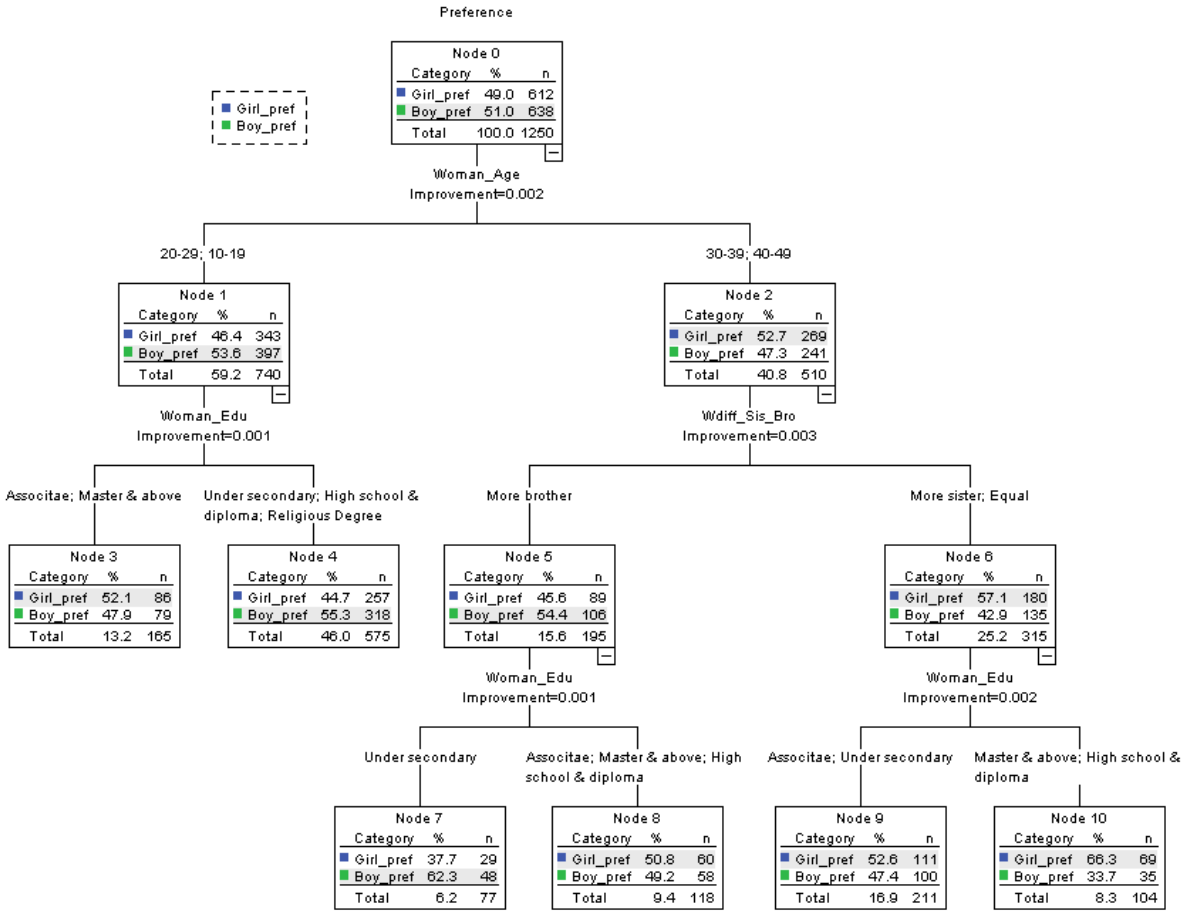

Figure 3. Decision Tree of Women's Sex Preference by Independent Variables

Misclassification matrix for classification model has been shown in Table (3) which indicates the accuracy of the classification model. The shaded cells in Table (3) signify correct classification or accuracy of the classification tree on Figure (3). The accuracy of the classification tree can be calculated in Equation (3). It shows the correct proportion of total number which predicted by tree. The results state that the accuracy of the model is 55 percent that indicates 55 percentages of women's sex preference have been classified correctly.

Risks and standard errors of the classification tree for training and learning data have been shown in Table (4). As mentioned before, to fit CART algorithm, data divided to two different groups of training and learning data and the model fits to these two groups. Indeed, training and learning data are used for fitting and confirming the validity of the model, respectively. When the risks of these two data sets are close to each other, it confirms the validity of the fitted model. According to the results of Table (4), these values are almost equal which indicates the validity of classification model proposed by the classification tree in Figure (3).

Table 3. Misclassification Matrix for Classification Tree

Observed Category	Predicted Category		Total
	Girl preference	Boy preference	
Girl preference	326	286	612
Boy preference	272	366	638
Total	598	652	1250

$$Accuracy = \frac{326+366}{1250} = 0.55 \quad (3)$$

Table 4. Risks and Standard Errors of the Classification Tree for Training and Learning Data

	risk	Standard error
Learning set	0.446	0.014
***k*-fold cross validity of training set**	0.483	0.014

4.Conclusions

Why should parents at all prefer having children of one sex over another one? It is generally argued that children of a particular sex are desired to provide certain utilities, such as financial, social, or psychological benefits. In developing countries, for example, sons are presumed to have greater economic net utility than daughters, since male offspring is better able to provide assistance in agriculture and fishing (Bhatia, 1984) and to serve as a simple social security system. In patrilineal societies sons are also valued for continuing the family name. Daughters, on the other hand, should be more reliable in providing old-age assistance. In addition, they are frequently desired to help with household tasks or to care for younger siblings. Thus, even in countries with a prevailing preference for sons, many families consider it important to have at least one daughter (Arnold, 1997).

Morgan et al. (1988), for example, have argued that boys might reduce parents' divorce risk, since fathers' attachments and obligations to their children, and marital cohesion, are greater if they have sons (Diekmann and Schmidheiny, 2004). Women, on the other hand, may consider girls as easier to be raised or as more rewarding companions (Marleau and Saucier, 2002).

Along the same lines, Brockmann (2001) argued that, for example, "increasing female labour-force participation and the growing 'burden of ageing' should increase the value of a daughter. Moreover, changing expectations concerning the division of labor and family responsibilities between men and women (Hank and Jürges 2007), followed by a more positive evaluation of women's role in society, might rather foster the development of a preference for girls (Hammer and McFerran 1988).

Due to the consequences of sex preferences for children on couple's fertility behavior such as sex-selective abortions and abnormally changes of sex ratios, it is important to investigate determinants of sex preferences. Most of researchers modeled sex preferences by logistic regression which was developed by statistician David Cox in 1958. It is a regression model where the dependent variable is categorical with two levels. Cases with more than two categories are referred to as multinomial logistic regression. Logistic regression measures the relationship between the categorical dependent variable and one or more independent variables by estimating probabilities using a logistic function, which has logistic distribution.
Hank and Kohler (2002) studied sex preferences by multinominal logistic regression according to predictors such as age, educational level, cultural variables and sex distribution of previous children in Germany. Rai et al. (2014) examined sex preferences by multinomial and Binary logistic regression in Nepal by selected predictors such as age, number of children, sex of the last child, educational level, and job and economic status. Frempong and Codjoe (2013) considered age, place of residence, region, religion, job status, lineage, and education, as predictors to study sex preferences in Ghanaians family by binary logistic regression. Mansurian and Khushnevis (2006) have considered age, education, the total number of children, marriage duration, place of residence, sex composition of children as nominated predictors and found their influence on fertility by applying logistic regression.
There are a number of reasons for difficulties of logistic regression to investigate sex preferences. First, this method is poorly suited for modeling many possible predictor variables. Second, interactions are generally difficult to model by this method. Third, logistic regression doesn't have efficient procedure to handle missing data. Due to the advantages of CART algorithm compared to logistic regression, in this article, CART algorithm was applied to 1250 women age 15-49 (kazemipour, 2014). Following results has been drawn from the extracted decision tree:

- Without considering any independent variables, women in this study preferred to have boys. This result is in favor of the results in countries such as India, China, and Korea. Pande and Astone (2007) concluded that this preference is deeply rooted in social, economic and cultural factors.
- Educated women in young (10-30) and old (30-49) age groups preferred to have girls. The same result has been announced by Jacobsen and Engholm (1999) between educated Swedish women. They mentioned that sometimes parents may desire a mixed sex because of the different benefits that accrue from each sex. Each partner, for example, might prefer to have at least one child of his or her own sex for the

purpose of companionship. Some authors like Wongboonsin and Ruffalo, (1995) mentioned that variations in sex preferences among countries and regions could be linked with factors involve the individual characteristics of parents, especially their level of education. The conclusion of Frempong and Codjoe (2013) was against girl preferences. They resulted that Ghanaian women who had higher education had higher likelihood to prefer son.

- Women's age categories have also an important role on the resulted decision tree. Sex preferences of women in young (10-30) and old (30-49) age groups were different. This is contrary to the findings of Westley and Choe (2007) in Pakistan. They resulted that young and old adults were less likely to prefer a daughter. Because sons are regarded as economic assets and security particularly during old age.
- Another important factor on women's sex preferences in this study was the difference number of women's siblings. Some of the researchers such as Lyngstad and Prskawetz (2010) considered number of siblings influences on fertility. They measured social interaction through the cross-sibling influences on fertility of Norwegian families. The authors studied the data which included the siblings' fertility, education, income, and marital histories. kazemipor (2014) also pointed to the influence of the number of women's siblings on their fertility. However, less attention has been devoted to the influence of women's siblings on their sex preferences for children.

References

Andersson, G., Hoem, J., and Duvander, A.Z. (2006). Social differentials in speed-premium effects in childbearing in Sweden. *Demographic Research*, 14 (4), 51-70, http://www.demographic-research.org/Volumes/Vol14/4.

Andersson, G. Hank, K., and Vikat, A. (2007). Understanding parental gender preferences in advanced societies: Lessons from Sweden and Finland, *Demographic Research,* Volume 17 Article 6, Pages 135-156, Published 12 October 2007, http://www.demographic-research.org/Volumes/Vol17/6/, DOI: 10.4054/DemRes.2007.17.6.

Arnold, F.(1992). Sex preference and its demographic and health implications. *International Family Planning Perspectives;* 18 (3), 93-101.

Arnold, F. (1997). 'Gender Preferences for Children'. *Demographic and Health Surveys Comparative Studies No.23*.

Arokiasamy, P.(2002). Gender Preference, Contraceptive Use and Fertility in India: Regional and Development Influences. *International Journal of Population Geography*. 8, 49-67.

Bhatia. J.C.(1984). Sex Preference of children in rural Ghana (West Africa). *Health and Population* - Perspectives & Issues 7(1), 32-47.

Bhukya D.P, and Ramachandram S. (2010). Decision tree induction-an approach for data classification using AVL–Tree. *Int J Comp d Electrical Engineering.* 2(4): 660-665. doi: http://dx.doi.org/10.7763/IJCEE.2010.V2.208.

Breiman L., Friedman J.H., Olshen R.A., Stone C.J. (1984). *Classification and Regression Trees*. Belmont, California: Wadsworth, Inc.

Brockmann, H. (2001): Girls preferred? Changing patterns of sex preferences in the two German states. *European Sociological Review*, 17 (2), 189-202.

Campbell, E.K., and Campbell, P. G. (1997). Family size and sex preferences and eventual fertility in Botswana. *Journal of Biosocial Science,* 29,191-204.

Chung, W., and Gupta, M.D. (2007). The decline of son preference in South Korea: The roles of development and public policy. *Population and Development Review*, 33(4):757–783.

Cox, D.R. (1958). "The regression analysis of binary sequences (with discussion)". *Journal of Royal Statistic Society*, B 20: 215–242.

Dahl, G.B. and Moretti, E. (2008). The Demand for Sons, *Review of Economic Studies*. 75, 1085–1120.

Dahl, E., Gupta, R.S. Beutel, M. Stoebel-Richter, Y. Brosig, B. Tinneberg, H.-R. and Jain, T. (2006). Preconception sex selection demand and preferences in the United States. *Fertility and Sterility*, 85 (2), 468-473.

Diekmann, A., and Schmidheiny, K. (2004). Do parents of girls have a higher risk of divorce? An eighteen-country study. *Journal of Marriage and Family*, 66 (3), 651-660.

Edlund, L., and Lee, Ch. (2013). Son Preference, Sex Selection and Economic Development: The Case Of South Korea, *Working Paper* 18679, http://www.nber.org/papers/w18679, national bureau of economic research, Cambridge, January 2013.

Francis, D.R.(2015). Do Fathers Prefer Sons? http://www.nber.org/digest/oct04/w10281.html.

Frempong, G.A., and Codjoe, S.N.A. (2013).Education and sex preferences of children in Ghana, Regional Institute for Population Studies, University of Ghana. P.O. Box LG 96. Legon, in XXVII IUSSP International Population Conference, Session 171: Poster Session on Fertility. Busan, Korea, 26-31 August 2013.

Fuse, (2010), Variations in attitudinal gender preferences for children across 50 less-developed countries, *Demographic Research*, http://www.demographic-research.org/Volumes/Vol23/36/. DOI: 10.4054/ DemRes. 2010.23.36.

Gelfand S. B., Ravishankar C. S., and Delp E. J.(1991).An iterative growing and pruning algorithm for classification tree design. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 13(2):163-174.

Gray, E., Morrison, N.M. (1974). Influence of Combinations of Sexes of Children on Family Size. *The Journal of Heredity*. 65: 169-174.

Gray, E.E., and Evans A. (2005). Parity progression in Australia: What role does the sex of existing children play? *Australian Journal of Social Issues*, 40 (4), 505-520.

Jacobsen, R., M_ller, H., and Engholm, G. (1999). Fertility rates in Denmark in relation to the sexes of preceding children in the family. *Human Reproduction*. 14: 1127-1130.

Hammer, M., and McFerran J. (1988): Preference for sex of child: A research update. *Individual Psychology*, 44 (4), 481-491.

Hank, k., Kohler, H.P. (2000). Gender Preferences for Children in Europe: Empirical Results from 17 FFS Countries, Demographic Research, Volume 2, Article 1, Published 15 January 2000, www.demographic-research.org/Volumes/Vol2/1/DOI: 10.4054/DemRes.2000.2.1.

Hank, K. and Kohler, H.P. (2002).Gender Preferences for Children Revisited: New Evidence from Germany, Mpidr Working Paper Wp ,017 MAY 2002.

Hank, K., and Jürges H. (2007). Gender and the division of household labor in older couples: A European perspective. *Journal of Family Issues*, 28 (3), 399-421.

Hejazi, N. (2013). Employed Women's attitude to second child and its inflectional factors, *Health System Research journal,* 9.

Kass, GV. (1980). An exploratory technique for investigating large quantities of categorical data. *Applied Statistics.* 29: 119-127.

Kazemipour, Sh.(2014). Childbearing Attitudes and its social, economical and cultural factors, Statistical Research Center, Tehran, Iran.

Kugler, A. D., and Kumar, S. (2015). Preference For Boys, Family Size And Educational Attainment In India Working Paper 21138 http://www.nber.org/papers/w21138 National Bureau Of Economic Research 1050 Massachusetts Avenue Cambridge, MA 02138 April 2015.12.06.

Lin, N., and Noe D, He X.(2006).Tree-based methods and their applications. In: PhamH, Ed. *Springer Handbook of Engineering Statistics*. London: Springer-Verlag; pp. 551-570**.**

Loh, W., andShih Y. (1997). Split selection methods for classification trees. *Statistica Sinica.***7**: 815-840.

Lundberg, S. (2005). Sons, daughters, and parental behavior. *Oxford Review of Economic Policy*, 21 (3), 340-356.

lyngstad, T.H., and Prskawetz, A. (2010). Do siblings' fertility decisions influence each other?, *Demography*, Volume 47-Number 4, November: 923–934

Mansurian, M.K., and Khushnevis, A. (2006). sex preferences of ever married women on fertility behaiviour; a case study in Tehran, *Social and human science of Shiraz journal*. 47.

Marleau, J.D., and J.-F. Saucier (2002): Preference for a first-born boy in western societies. *Journal of Biosocial Science*, 34 (1), 13-27.

McDougall, J., David J., DeWit, G., and Edward E. (1999). Parental Preferences for Sex of Children in Canada*, Sex Roles, Vol. 41, Nos. 7/8.*

Mishra, V., Roy, T. K., and Retherford, R. D. (2004). Sex Differentials in Childhood Feeding, Health Care, and Nutritional Status in India. *Population and Development Review* 30, 269- 295.

Morgan, S.P., Lye, D., and Condran G. (1988): Sons, daughters and the risk of marital disruption. *American Journal of Sociology*, 94 (1), 110-129.

Pande, Rohini P., and Astone, N.M. (2007). "Explaining son preference in rural India: the independent role of structural versus individual factors." *Population Research and Policy Review* 26 (1):1-29.

Park, C.B., Cho, N.H. (1995). Consequences of Son Preference in a Low-fertility Society: Imbalance of the Sex Ratio at Birth in Korea'. *Population and Development Review*. 21: 59-84.

Pollard, M., Morgan S.P. (2002): Emerging parental gender indifference? Sex composition of children and the third birth. *American Sociological Review*, 67 (4), 600-613.

Quinlan, R.J. (1993). *C4.5: Programs for Machine Learning*. San Mateo, California: Morgan Kaufmann Publishers, Inc.

Rai, P., Paudel, I.Sh., Ghimire, A., Pokharel, P.K., Rijal, R. and Niraula, S.R. (2014). Effect of gender preference on fertility: cross-sectional study among women of Tharu community from rural area of eastern region of Nepal. *Reprod Health*. 11: 15. doi: 10.1186/1742-4755-11-15.

Raley, S., and Bianchi S. (2006). Sons, daughters, and family processes: Does gender of children matter? *Annual Review of Sociology,* 32, 401-421.

Scharping, T. (2003). *Birth Control in China, 1949-2000: Population Policy and Demographic Development*. Routledge, London and New York, 2003.

Seidl, C. (1995).The desire for a son is the father of many daughters. A sex ratio paradox. *Journal of Population Economics*, 8 (2), 185-203.

Shahbaziyan, S., Gholami, A., Shahbaziyan, S. (2014). The Rule of Sex Preferences in Fertility Behavior of Women in Kangavar, *Journal of Ilam medical science*, 22(6).

Timofeev, R. (2004), *Classification and Regression Trees (CART)Theory and Applications*, A Master thesis, CASE - Center of Applied Statistics and Economics Humboldt University, Berlin.

Westley, S.B., and Choe, M.K. (2007). *How does son preference affect population in Asia? Analysis from the East-West*, center. No.84.

Wongboonsin, K and Ruffolo, V, P. (1995). Sex preference for children in Thailand and some other South-East Asian Countries. *Asia-Pacific Population Journal* 10(3), 43-62.

Zhu, W.X., Lu,L. and Hesketh. Th. (2009). China's excess males, sex selective abortion, and one child policy: analysis of data from 2005 national intercensus survey. *British Medical Journal*, 338.