Measuring cost of children using equivalence scale on Japanese panel data

Critical evaluation of a published article 25.11.2014

Motivation

Fertility rate is currently a big problem all over the world. Undeveloped countries have too high fertility and then cannot use their new labor force sufficiently because people cannot get jobs there. Their goal is to decrease fertility rate on controllable level. Developed countries have different problems. They have low fertility rates and a problem with a lack of labor force. Many developed countries try to solve this problem with immigration but that only partly resolves the problem of low birth rate. In order to increase the birth rate, developed countries tried different approaches, which try to decrease the cost of raising children. Those measures decreased cost of bearing a child slightly, but not enough. Fertility rates in most developed countries are still low and Japan is one of the countries where fertility rate is a big problem. In order to analyze the reasons for low fertility rate, many studies have been done and this is one of them. Japan has small birth rate even compared to some other developed countries and author therefore tries to analyze what are the costs of bearing a child in Japan so this would give a good indicator for policy measure of what can be improved. There was a hypothesis that fertility decline in Japan happened because of high cost of child rearing. The author analyzes costs of rearing children in Japan during the years 1993 and 1999.

Summary of the paper

The study is examining why the total fertility rate in Japan has been in decline. Total fertility rate is the average number of children that are born to a woman over her lifetime. The delay of childbearing among young married couples is said to be one of the most important reasons for decline in fertility rate in Japan. The latter may be because of the greater expenditure on children's education. Therefore in this research the cost of child is estimated using an equivalence scale. The results are then compared with the results of other researches in Japan and other developed countries.

Equivalence scale predicts that household utility is considered under budget constraint $p'q = y_i$, where p is the price vector ($p \in R^n$), *y* is the household income and *q* is the commodity vector. Utility function is defined as $u(q, z)$. From that we get household cost function *C(u,p,z)*. Here *z* is the household composition. It is assumed in the model that households with same composition have an identical preference which leads to identical household cost function. The author also provided the table from National fertility survey in 2002 with reasons for having fewer children than ideal. In his research he tried to take them into account. There are many methods to measure the cost of children and this paper uses Rothbarth model and income satisfaction measure.

Estimation using Rothbarth model:

In this model we have two groups of the goods and services consumed by a household. One is the 'adult goods,' which are consumed exclusively by the adults (husband and wife) for example alcohol or adult clothing. The other one is total expenditure, which are goods and services consumed jointly by adults and children. In his research, author got the data for those variables from the Household Survey by the Institute for the Research on Household Economics. These are the data of Japanese Survey on Consumers.

Estimation using 'satisfaction with income:'

In this estimation, it is measured how couple feels about their income, that is the relative satisfaction is used to present couple's welfare level. In this model, a general utility function is allowed since welfare is directly measured. The author used variable 'IncomeSatisf,' which takes discrete values from 1 to 4 where 4 means highest satisfaction. The other variable is total expenditure, which was explained before. The data comes from the same source, but variable IncomeSatisf was only obtained from 1995 to 1999.

The author then estimated the two models with following regressions:

$$\text{ExpenHusWi} = \alpha_0 + \alpha_1 \text{ LnExpTotal} + \sum_j \alpha_j n_j + \sum_k \alpha_k YearDummy_k + u_1$$

for Rothbarth model and

$$\text{IncomeSatisf} = \alpha_0 + \alpha_1 \text{ LnExpTotal} + \sum_j \alpha_j n_j + \sum_k \alpha_k YearDummy_k + u_2$$

for estimation using satisfaction with income. Here $n_j$is the number of children in each age category. YearDummy is 1, if it certain event was observed in year k. In this way he controlled for the effects of time. ExpenHusWi is monthly total expenditure for husband and wife. Equivalence scale is computed as $\exp(-\frac{\alpha_j}{\alpha_1})$. In its research author used only 1 alpha and allowed n to be discrete numbers (0,1,2,…). The author observed equivalence scale for children aged 0-18 and for children aged 0-6, 7-13 and 14-18.

The estimates obtained for Rothbarth model show that equivalence scale is around 1.12. It was regressed using pooled OLS estimator, fixed effects estimator and random effects estimator. Author also made a Hausman test in which he compared FE with RE and did not reject the null hypothesis and therefore concludes random effects estimator performs better. Also Breusch-Pagan test was performed where he compared pooled OLS with random effects and also concluded that random effects estimator performs better in this situation. So he took random effects estimate of equivalence scale as given. That yielded results for cost of bearing a child of 1.13 for a child between age 0 and 18 and 1.124 for child at age 0-6, 1.126 for a child aged 7-13 and 1.26 for a child aged 14-18. R-square statistics were around 0.24 for different tests. These results were then compared with the results of other researches and author found out that they obtained similar results. Also for each test (pooled OLS, fixed and random effects) two test were made, one for the cost of child aged 0-18 and the other one for cost of child aged in three different categories.

The estimates obtained using satisfaction with income on the other hand show different equivalence scale. The estimations were computed under pooled OLS, ordered probit, random effect ordered probit and fixed effect regressions. In this case no tests were made which is better, which is in part because it is hard to compare them, but all estimates were reported. Estimates of equivalence scale were between 1.386 and 1.454 for the cost of bearing a child aged 0-18 years. Also similar estimates are for the child of each age group where there were the biggest estimates for equivalence scale estimate for child in age

group 14-18 years. Here the results of equivalence scale are between 2.090 (for ordered probit estimation) and 4.329 (fixed effects estimation).

Based on that results the conclusion was that using traditional Rothbarth model, the cost of children is not high, while estimation results on new subjective scale approach showed that cost of children is significantly higher in Japan then in other countries. The author proposed there should be a further research performed in that subject in future.

Good and not so good things about the paper

Things I do not like so much about the paper

In the second research there were four different estimators, namely pooled OLS, ordered probit, random effect ordered probit and fixed effect. In this way one cannot test estimates using Hausman test to decide which one is the best. Author then just interprets the results as a whole because all tests give fairly similar results. I think in this case author should include likelihood estimates and compare at least ordered probit and random effect probit. Also fixed effect and pooled OLS could be compared. Equivalence scales do not differ a lot in different models (that is why author made conclusions by saying they are higher than in Rothbart model), but their t-statistics and significance does. That is why I think author should compare at least what is comparable and interpret results with care, because it cannot determine which estimator is the best. Also I think author should make fixed effect ordered probit estimator (although it could be hard to implement) and compare it with random effect ordered probit with one of the tests (preferably Hausman test). More comparison between models would be useful because results are not exactly the same, especially if we look at the significance of different estimators.

Next thing that I think is not so good is that author should include confidence intervals in calculating estimates for the equivalence scale – also it should give rough estimates for equivalence scale about its confidence intervals. This is especially important in estimation using satisfaction with income, because some estimates are pretty low (and in some cases they have also low t-statistics) which could lead to wrongly computed equivalence scale. Equivalence scale is computed as $\exp(-\frac{\alpha_j}{\alpha_1})$. So confidence intervals for α_j and α_1 should be computed as well as the new estimates for equivalence scale which include confidence intervals. The differences could be high – for example for the fixed effects estimator α_j which represents child at ages between 0-18 cost is not significant, which consequently means that equivalence scale in lower 95% confidence interval will be lower than one – which is much less than 1.427 as estimated by author. Also results in other tests it will yield broader results. Inclusion of confidence interval might change the final results.

I tried to calculate confidence intervals myself for Child0-18 just as an example what could be done. Because equivalence scale depends always on 2 values, I always took one as given and looked at the 95% confidence intervals of the other;

95% confidence interval for	Pooled OLS	Fixed effect	Random effect

lnExpTotal lower interval	1,13	1,11	1,12
lnExpTotal upper interval	1,14	1,12	1,13
Child018 lower interval	1,11	1,05	1,1
Child018 upper interval	1,16	1,16	1,15

Table 1: Confidence intervals for equivalence scale in Rothbarth model; results do not change substantionally, estimates for random effect estimator are between 1,1 and 1,15

	Pooled OLS	Ordered probit	Random effect ordered probit	Fixed effect
lnExpTotal lower interval	1,27	1,27	1,20	0,79
lnExpTotal upper interval	1,53	1,53	1,85	2,68
Child018 lower interval	1,32	1,32	1,36	1,27
Child018 upper interval	1,45	1,45	1,67	2,63

Table 2: Table 1: Confidence intervals for equivalence scale in life satisfaction model; here results for different estimators differ quite a lot; another reason to test which estimator performs better (where it could be tested).

There is also one thing that I do not like that much, but I understand that it is hard to fix that. In Rothbart model individuals were observed between years 1993 and 1999, but in the second model (satisfaction with income) only individuals between years 1995 and 1999 were observed. The author controlled the time series differences with year dummies. But in different time there could be different laws already passed and higher or smaller problems with fertility in the country than in other years. So results might be different because of that. Also author made a comparison between those two models although they are compared in different time frame which is not necessarily good.

Another thing that is worth of a mention is the results in estimation with satisfaction with income. As argued before, at fixed effects estimator, the estimate for children aged 0-18 is not significant (one could argue the same for child aged 0-6 and 7-13). Another anomaly is at pooled OLS. Beta estimate for variable Child018 is positive, meaning that equivalence scale for them is calculated wrong (should be lower than 1). Also when the estimation for three different generations of children is done, estimates for their variables become negative. There are significant changes in all variable estimators excluding log of monthly total expenditure. And in both cases both different estimators are significant. Another thing here is that when estimating coefficients he used different number of observations (in pooled OLS for satisfaction with income) – 5335 when estimating Child018 and only 5183 when estimating Child06, Child713 and Child1418. The author did not explain why he used different number of observations. I think pooled OLS is also not the best estimator for such time series and for this estimation. At the lectures we have shown that it is usually not efficient in time series. It was also shown that in cases when we are doing time series analysis like this one, pooled OLS might not be the best estimator.

There could be also an improvement of the paper if author posted number of observations where Child06, Child713 and Child1418 variables were different than zero. Because Japan is dealing with low fertility and author used smaller number of samples in satisfaction with income, it would be good to know number of observations in those categories. The reason why that that might be a problem is that author states that data consists of married women aged between 24 and 34 (in year 1993). That means that the oldest women were in 1999 aged 40 years. Because average age of first child bearing in developed countries is low and a big majority of the women in the sample was below 40 years old (also many new observations with younger women were added in 1997), it could happen that there were not

many observations of women who had one or more child aged 14-18. And because those estimators are based on asymptotic properties, different statistics and estimates could be estimated wrong.

Thing I like about paper

This paper consists of two estimations. In first estimation (Rothbart model) the author performed Hausman test for comparing fixed effects estimator with random effects estimator. The test showed that random effects estimator is better. Furthermore there is also Breusch-Pagan test where author tested pooled OLS estimator with random effects estimator. The test showed that random effects estimator is better. On the basis of these two tests, author therefore concluded that random effects estimator is the best estimator. Also results of the other two tests are shown and they do not differ much from the random effects estimator. So because of this I think it is good to take values obtained by random effects estimator as the best estimates and author did that. I think it is a good thing that author made a test to determine which model is the best.

Another good thing is that author included time dummies which showed to be significant for some years in both models with different estimators. In this way author successfully controlled time effects, for example that people in year 1998 were in general less satisfied with income. If he did not do this, those differences in years could have had bad effect on overall results.

It is also good that author included different generations of children. By this I mean inclusion of the variables which measure number of children between age 0 and 6, 7 and 13 and 14 and 18. The estimates for equivalence scale for children aged between 14 and 18 all seem significant and there seems to be high equivalence scale, and also pretty significant. They also show at which ages of the children is the equivalence scale the highest and by doing that its final results are more informative.

I also think that in estimation with satisfaction with income the author did a good choice including ordered probit and random effect ordered probit estimators. That is because it is dealing with discrete positive values of dependent variable. Probit models seem to be the good decision there. Also when including variables child of certain age group, he allowed different discrete numbers to be in such variable. He could make just a dummy there, whether the family has a child in certain age group, but that would not give the best estimates because families with more than one children in certain category would have higher expenses for children at that age.

Another good thing about that paper is the selected data. With selected data the author successfully controlled reasons for having fewer children than ideal. He chose same individuals (and he chose his target group – married young women) and he observed them during seven years. That period is good because that is when women in research were most fertile so the results from the model should therefore show quite good picture of reality.