

Econ 7100

## Final Project Evaluation Report

Submitted by Group#3

Jingxian Liu, Qiang Deng, Yang Bo, Yuchong He, Zihan Wang,

April 14th 2020

### **Introduction**

#### Background and Overall Purpose

Many first-time mothers living in poverty in Nashville have a hard time taking good care of their babies. A randomized intervention of providing relative information through giving out baby-books is implemented aiming to improve the mothers' understanding on how to take care of their babies. This evaluation will identify the intended outcome of total mother's knowledge (including nutrition, safety, parenting and development knowledge that the mothers could possibly gain from reading the given books), and examine the impact of this randomized intervention on this outcome. In addition, the evaluation will come up with conclusions and recommendations.

#### Study Design Overview

In order to measure the effect of this intervention, total of 198 first time mothers were randomly assigned into three groups:

- 1) an educational book group (the intervention)
- 2) a non-educational book group
- 3) a no book group

A questionnaire designed to collect mothers' baseline characteristics was first distributed before any intervention. In addition, mother was also required to complete an exam consisting 31 questions on how to take care of babies under different situations. Then in 2, 6, 9, 12, 18 months after the baby is born, a follow-up survey on mother's background, as well as another exam was distributed to each of member of the three groups. We collect the data in the process, and we use the percentage of corrected answers as the indicator on how the book impact mother's knowledge on taking care of babies. The outcomes of the intervention will be discussed in the Main Results part of the report.

## Methodology

### Survey Approach

The survey used mainly quantitative data collection method. Participants self-reported all the outcomes of interest on the interviews, that were given either by phone (12 times) or in person (7 times). Data collection consisted primarily of closed ended questions and respondents generally selected responses from a set of options.

### Data Analysis Procedures

Data was originally collected using the questionnaires. We then tailored the dataset towards our analysis, and these data outputs are used to generate the narrative of this report.

First of all we generated a variable in order to measure our outcome of interest, mom's knowledge overtime. We created a variable "KKnowledge\_Baseline" denoting the percentage of correct answers among all the questions about overall knowledge at baseline, and through testing, assured AW01BPC1 to be our outcome.

Then by testing the relevance of other variables to the outcome of interest, we included college\_degree, good\_health, paid\_job, and two dummy variables treatment1v2 (0 if in group2, 1 if in group1) and treatment1v3 (0 if in group3, 1 if in group1).

By stacking up the regression, the variables we finally included are *StudyGrp AWFamID AWBAgeD AWBAgeD2 total\_knowledge WaveSort treatment1v2 treatment1v3 AWBQ02E1 AWBQ02E2 AWBQ02E3 AWBQ02E4 AWBQ02E5 AWBQ03* (these are indicators on mother's racial status) *college\_degree good\_health paid\_job AWBQ07*(whether the mother is living alone) *AWBQ17*(family income) *AWBQ20* (whether the pregnancy is planned) and *AWBQ34*(part of any book-giveaway program). Thus, we reshaped the original dataset into a panel dataset on the outcome of moms' total knowledge.

## Establishing Counterfactuals

To begin with, we perform the following test to compare characteristics between the treatment group, namely group1 and control groups, namely group2 and group3 at baseline. As randomized assignment of treatment requires, we need to have an “identical” control group in order to correctly estimate the impact of intervention. To validate this assumption, we test whether eligible mothers in the treatment and comparison groups have similar characteristics at baseline, as shown in table 1.

Convariates	Characteristics	Group 1v2			Group 1v3		
		Difference	t-stat	Significance	Difference	t-stat	Significance
AWBQ02E1	African American	-0.0733138	-0.8786	No	-0.1525346	-1.8193	No
AWBQ02E2	Amer. Ind./AK Native	N/A	N/A	N/A	N/A	N/A	N/A
AWBQ02E3	Asian	0.0151515	0.0146628	No	0.0142857	0.9407	No
AWBQ02E4	Caucasian/white	0.0146628	0.1863	No	0.1419355	1.733	No
AWBQ02E5	pacific Islander	N/A	N/A	N/A	N/A	N/A	N/A
AWBQ03	Hispanic/Spanish	0.0404868	0.8061	No	0.0058548	0.1311	No
college_degree	college_degree	-0.0229717	-0.4009	No	0.0995392	1.4809	No
good_health	good_health	0.0190616	0.4707	No	-0.0069124	-0.1561	No
paid_job	paid_job	0.0712866	0.7984	No	0.0838407	0.9534	No
AWBQ07	live alone	-0.1197458	-1.8381	No	-0.0640417	-0.9213	No
AWBQ17	income last year	-0.0986928	-0.1776	No	1.382656	2.3776	Yes
AWBQ20	planned pregnancy	0.0537634	0.8702	No	0.2013825	2.853	Yes
AWBQ3	Book giveaway program	-0.0342131	-0.911	No	-0.0502304	-1.5605304	No

The conclusions is that there is no significant differences between group 1 and group 2 in the selected observed characteristics at baseline. However, comparing group1 and group3, we have two characteristics (income from last year and whether the mothers are planned pregnant) that are distinct from 0. The money made last year was measured on a scale from 1 to 9, with 1 is equivalent to \$4,000 increment in annual income.

Consequently, compared to group 3, group 1 mothers' families on average made ~\$5,520 more compared to those from group 3. The other difference lies in the part on whether the pregnancy is planned. On average, mother from group 1 is almost 20% more likely to have a baby she has planned for. These two differences undermine the validity of group 3 as a perfect counterfactual.

## Main Results

In examining the impact of the intervention, we used pre-post analysis, randomized assignment analysis, and difference-in-difference analysis.

### Pre-Post Analysis

First, we perform several t-tests against total mother's knowledge by different groups.

- Difference in total knowledge of Studygroup1

WaveSort	Mean	Std. Err.	T-Score
1	39.07355	1.256905	
4	52.63636	2.964023	
Diff	-13.56282	3.093142	-4.3848

According to the result, mothers in the first treated group have their total knowledge increased by 13.56 over the first 6 months. The corresponding t-score is -4.38, which makes the difference significantly different from 0.

WaveSort	Mean	Std. Err.	T-Score
4	52.63636	2.964023	
6	56.7054	2.821156	
Diff	-4.069036	4.111066	-0.9898

As we can see in the chart, mothers in the first group have their knowledge rose by 4.07 over the second 6-month period. The T-score is -0.9898, which means the difference is insignificantly different from 0. Mothers' total knowledge not increase obviously in the second 6-month period.

WaveSort	Mean	Std. Err.	T-Score
6	56.7054	2.821156	
7	63.46227	2.741264	
Diff	-6.756874	3.957146	-1.7075

As for the third 6-month period, mothers' total knowledge has increased by 6.76. The T-score is -1.7075, which is bigger than -1.96, but smaller than -1.65. The difference is insignificantly different from 0 at 95% confidence level, but it is statistically different from 0 at 90% confidence level.

- Difference in total knowledge of Studygroup2

WaveSort	Mean	Std. Err.	T-Score
1	39.05227	1.269362	
4	40.17544	2.684206	
Diff	-1.123166	2.842911	-0.3951

Mothers in the study group 2 have their knowledge increased by 1.12. The T-score is -0.395, which means the difference is insignificantly different from 0. The total knowledge not increased obviously in the first 6 months in the second group.

WaveSort	Mean	Std. Err.	T-Score
4	40.17544	2.684206	
6	50.56577	2.601054	
Diff	-10.39033	3.751092	-2.7699

Over the second 6-month period, the total knowledge of mothers in study group 2 has increased by 10.39. The T-score is -2.77, which means the difference is significantly different from 0. The total knowledge increased evidently in this period in this group.

WaveSort	Mean	Std. Err.	T-Score
6	50.56577	2.601054	
7	53.35575	2.850176	
Diff	-2.789976	3.850582	-0.7246

The outcome increases by 2.79 over the third 6-month period. The corresponding T-score is -0.7246, which makes difference insignificantly different from 0. The outcome has not changed apparently in this period in study group 2.

- Difference in total knowledge of Studygroup3:

WaveSort	Mean	Std. Err.	T-Score
1	42.02771	1.264332	
4	48.18966	2.319394	
Diff	-6.161941	2.526251	-2.4392

In study group 3, the total mothers' knowledge has increased by -6.16 over the first half year. The corresponding T-score is -2.439, which means the difference is significantly different from 0 at 95% confidence level.

WaveSort	Mean	Std. Err.	T-Score
4	48.18966	2.319394	
6	55.31439	2.621048	
Diff	-7.12473	3.496455	-2.0377

Mothers total knowledge has increased by 7.12 over the second half-year, the corresponding T-score is -2.04, which means the difference is significantly different from 0 at 95% confidence level.

WaveSort	Mean	Std. Err.	T-Score
6	55.31439	2.621048	
7	55.12778	1.86104	
Diff	0.1866071	3.245699	0.0575

Total knowledge has increased by 0.187 over the third 6-month period, the corresponding T-score is 0.0575, which is smaller than 1.96. The difference is insignificantly different from 0. Total mothers' knowledge changed little in this period.

According to the several t tests, we could see that the mothers in the study group 1 have their knowledge increased over the whole evaluation period, especially during the 3<sup>rd</sup> 6-month period.

Although mothers in the study group 2 and study group 3 have their knowledge increased throughout the whole period, the difference in both groups are smaller than the study group 1's. Thus, we could infer that the program could help mothers increase their total knowledge, and the impact will be increasingly significant as program goes on.

Then, we use simple regression and multivariate regression to figure out what factors have contribution to the increase of total mothers' knowledge in different groups.

- Regression against total knowledge of study group 1:

Simple regression:

total knowledge	Coef.	Std. Err.	t	P> t
round	7.885081	0.8511415	9.26	0.00
_cons	41.14764	1.405021	29.29	0.00

According to the chart, we could find the coefficient of round is about 7.88, the corresponding p value is smaller than 0.05, thus we could say that the coefficient is different from 0. This means, mothers who receive the parental books would have their total knowledge increased by 7.885 every 6 months. The \_cons term states that in round 0 (baseline), the average of total knowledge about the treated group is 41.15.

#### Multi-Variate Regression:

total_knowledge	Coef.	Std. Err.	t	P> t
round	9.671501	1.324495	7.30	0.000
AWBQ02E1	-4.560762	2.286376	-1.99	0.052
AWBQ02E2	0	(omitted)		
AWBQ02E3	0	(omitted)		
AWBQ02E4	0.0690789	2.364829	0.03	0.977
AWBQ02E5	0	(omitted)		
AWBQ03	5.361392	1.78619	3	0.004
college_degree	3.053128	2.766671	1.1	0.275
good_health	4.99124	3.350233	1.49	0.143
paid_job	4.025415	2.465343	1.63	0.109
AWBQ17	1.423864	0.5008901	2.84	0.006
AWBQ34	1.182921	4.082552	0.29	0.773
_cons	31.30397	3.934186	7.96	0.000

As we include more variables in the regression, the coefficient of round is 9.67, and the P value is smaller than 0.05, which means the coefficient is statistically different from 0 at 95% confident interval. Mothers in the treated group would increase their total knowledge by 9.67 in every 6 months. The P value of AWBQ02E1, AWBQ03, AWBQ17 is smaller than 0.1. We could say that the coefficients of these variables are statistically different from 0 at 90% confidence interval. The coefficient of AWBQ02E1 is -4.560762, which means that black people will have 4.56 lower percentage to answer the questionnaire correctly compared to others who are not. And mothers who are Hispanic have 5.36% higher to answer the questionnaire correctly compared to the other mothers who are not. The increase in income by 4000 dollars would make the outcome higher by 1.42.

The \_cons term's coefficient is 31.3, which means the average score of the total knowledge in this group is 31.3 at baseline.

Other covariates' P value is bigger than 0.1, making their impact statistically insignificant.

- Regression against total knowledge of study group 2:

#### Simple regression:

total_knowledge	Coef.	Std. Err.	t	P> t
round	5.275269	0.8583215	6.15	0.00
cons	37.88945	1.452844	26.08	0.00

The coefficient of round is 5.28, which means mothers who receive random books have their knowledge increased by 5.28 every 6 months. The corresponding p value is smaller than 0.05, thus we could say that the coefficient is different from 0. The \_cons term means total mothers' knowledge in study group 2 is about 37.89 at baseline.

Multi-Variate Regression:

total_knowledge	Coef.	Std. Err.	t	P> t
round	5.091765	1.767388	2.88	0.006
AWBQ02E1	3.773961	4.075012	0.93	0.359
AWBQ02E2	0	(omitted)		
AWBQ02E3	29.23936	4.377456	6.68	0.000
AWBQ02E4	16.42338	3.367288	4.88	0.000
AWBQ02E5	0	(omitted)		
AWBQ03	9.064519	3.204535	2.83	0.007
college_degree	8.805175	3.851084	2.29	0.027
good_health	-10.36629	6.163141	-1.68	0.099
paid_job	5.944431	2.607497	2.28	0.027
AWBQ17	-0.642744	0.7070727	-0.91	0.368
AWBQ34	3.904919	5.845188	0.67	0.507
_cons	40.24078	6.88313	5.85	0.000

The coefficient of round is 5.0, and the P value is smaller than 0.05, making the difference statistically different from 0. We could say that mothers' total knowledge would increase by 5.09 in every 6 months. The P values of covariates AWBQ02E3, AWBQ02E4, AWBQ03, college\_degree, good\_health, paid\_job are smaller than 0.1, which makes the coefficient significantly different from 0 at 90% confidence level. According to the result, we could state that Asian mother has 29.24% higher to answer the questionnaire correctly than the mother who is not, and Caucasian mother has 16.423% higher to answer the questionnaire correctly than the mother who is not. Hispanic mothers have 9.06% higher accuracy than other mothers who are not. Mothers who have college degree would have 8.8% higher accuracy than other people. Mothers who are in good health will have 10.37 % lower accuracy compared to others who are not. And mothers who have a paid job would have 5.9% higher accuracy than others.

The \_cons term's coefficient 40.24, which means the average score of the total knowledge in this group is 40.24 at baseline.

Other covariates' P value is bigger than 0.1, which makes their impact statistically insignificant.

- Regression against total knowledge of study group 3:

Simple regression:



total_knowledge	Coef.	Std. Err.	t	P> t
round	4.743396	0.6969031	6.81	0.00
_cons	43.02287	1.322194	32.54	0.00

The coefficient of round is 4.74 in the study group 3. The P value is smaller than 0.05, which means the coefficient is significantly different from 0. Thus, we could say that mothers who do not receive books will have their knowledge increased by 4.74 every 6 months. The average of total knowledge about the study group 3 is 43.023 at baseline.

#### Multi-Variate Regression:

total_knowledge	Coef.	Std. Err.	t	P> t
round	5.707003	2.273846	2.51	0.015
AWBQ02E1	-13.14426	4.356393	-3.02	0.004
AWBQ02E2	0	(omitted)		
AWBQ02E3	-14.35532	5.260292	-2.73	0.009
AWBQ02E4	-8.141242	3.601797	-2.26	0.028
AWBQ02E5	0	(omitted)		
AWBQ03	-4.068672	3.991136	-1.02	0.313
college_degree	16.60528	4.439218	3.74	0.000
good_health	2.176646	3.929813	0.55	0.582
paid_job	-1.180502	3.434697	-0.34	0.733
AWBQ17	-1.371616	0.643489	-2.13	0.038
AWBQ34	8.819618	5.425629	1.63	0.110
_cons	54.81241	5.977198	9.17	0.000

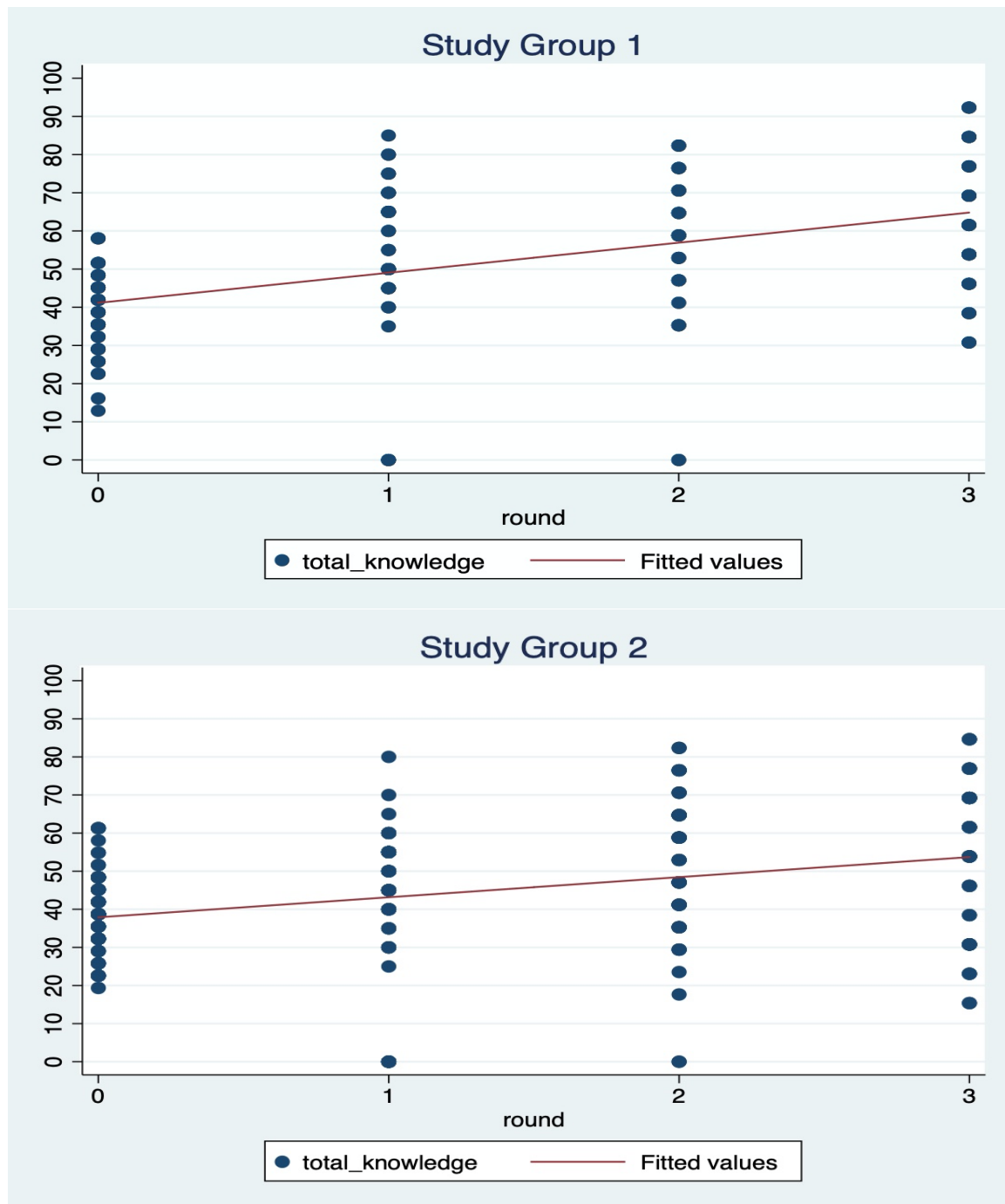
The coefficient of round is 5.71. The P value is 0.015, which means the coefficient is significantly different from 0. Mothers who do not receive books have their scores increased by 5.71 in every 6 months. The P values of covariates AWBQ02E1, AWBQ02E3, AWBQ02E4, college\_degree, AWBQ17 are smaller than 0.1, which means the coefficient is significantly different from 0 at 90% confidence level. According to the chart, we could know black mothers would have 13.14% lower accuracy compared to the other mothers. Asian mother has 14.35% lower accuracy than the mother who is not, and Caucasian mother has 8.141242% lower accuracy than the mother who is not. Mothers who have a college degree would have 16.6% higher accuracy compared to others who do not. And the increase in income by \$4000 would let the accuracy decrease by 1.37.

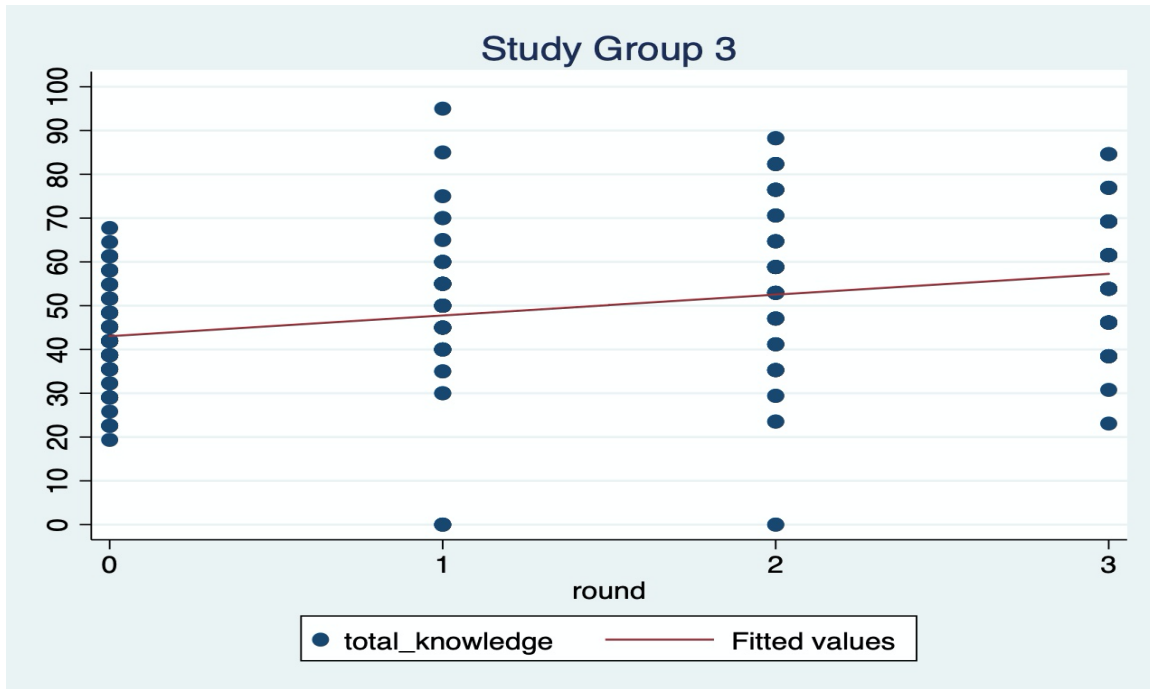
The \_cons term's coefficient is 54.81, which means the average score of the total knowledge in this group is 54.81 at baseline.

Other covariates' P value is bigger than 0.1, which makes their impact statistically insignificant.

Although the covariates show different influence on the outcomes in three groups, we could also see that mothers who read parental books have their knowledge increased significantly compared to the other two groups. Mothers read random books might not have significant help in total knowledge increase according to the similar coefficient of round in group 2 and group 3.

We could also see the rate of change through graphs.





As we can see in three graphs, three groups all have their outcome improved over the period of study. The rate of change is probably linear, as the straight line fits the rate of change. Group 1, whose line is the steepest, has the most significant increase in the outcome. Group 3 has the flattest line, which means that this group has the least increase in the outcome. These three graphs correspond to the estimation we did above.

### Randomized Assignment Analysis

Then we use randomized assignment analysis to evaluate the impact of the intervention on mothers' total knowledge, which is the quantified outcome through the correct percentage mothers answer in the questionnaire.

Then, we perform the second check to estimate the ITT effect. To achieve this, we compare Effect of intervention on at follow-up waves. In the first place, we analyze the ITT effect of treatment1v2. With the validity of the comparison group established, we can estimate the counterfactual as the average total knowledge of eligible mothers in the different groups. Table 2 shows the average mothers' total knowledge for eligible mothers in the treatment and comparison groups (Group1v2). Notably at baseline, the average total knowledge in the treatment and comparison groups are not statistically different, as should be expected under randomized assignment. Table 3 shows the average mothers' total knowledge for eligible mothers in the treatment and comparison groups (Group1v3).

Table 2 Evaluating ITT: Randomized Assignment with Comparison of Means (1v2)

	<b>Group1-Tx</b>	<b>Group 2</b>	<b>Difference</b>	<b>t-stat</b>
Mothers' total knowledge at baseline	39.07355	39.05227	-0.0212758	-0.0119
Mothers' total knowledge at follow-up (Wave=4)	52.63636	40.17544	-12.46093	-3.1207
Mothers' total knowledge at follow-up (Wave=6)	56.7054	50.56577	-6.13963	-1.602
Mothers' total knowledge at follow-up (Wave=7)	63.46227	53.35575	-10.10653	-2.5496

Table 3 Evaluating ITT: Randomized Assignment with Comparison of Means (1v3)

	<b>Group1-Tx</b>	<b>Group 3 (no book)</b>	<b>Difference</b>	<b>t-stat</b>
Mothers' total knowledge at baseline	39.07355	42.02771	2.954166	1.6503
Mothers' total knowledge at follow-up (Wave=4)	52.63636	48.18966	-4.446708	-1.1883
Mothers' total knowledge at follow-up (Wave=6)	56.7054	55.31439	-1.391014	-0.3614
Mothers' total knowledge at follow-up (Wave=7)	63.46227	55.12778	-8.334495	-2.5885

Given that we have a valid counterfactual of group 2, we can find the impact of the intervention simply by taking the difference between the average total knowledge of mothers in the treatment groups and randomly assigned comparison groups in the follow-up period. The impact is an increase in score of 12.46093 over 6 months. Replicating this result through a linear regression analysis yields the same result, as shown in **Table 4**. Finally, we run a multivariate regression analysis that controls for some other observable characteristics of the sample groups, and we find that the program has improved the total knowledge of mothers by 11.29863 over 6 months, which is nearly identical to the linear regression result.

As a result, for the comparison between group1 and group 2, we can be confident that no factors are present that are systematically different between the treatment and comparison groups that might also explain the difference in mothers' total knowledge. Both sets of

groups started off with very similar average characteristics and have been exposed to the same set of backgrounds and programs during the several six months of treatment. Then the only plausible reason that mothers' total knowledge increase over the treatment period is that group1 received the educational books and group2 received non-educational books.

Table 4 Evaluating ITT effect: Randomized Assignment with Regression Analysis (1v2)

	Linear regression	Multivariate linear regression
Estimated impact on mothers' total knowledge (Wave=4)	12.46093 (3.992958)	11.29863 (2.369822)
Estimated impact on mothers' total knowledge (Wave=6)	6.13963 (3.832565)	8.371659 (2.936751)
Estimated impact on mothers' total knowledge (Wave=7)	10.10653 (3.963974)	10.89637 (4.092797)

As shown in Table 5, a linear regression analysis yields the result.

Table 5 Evaluating ITT effect: Randomized Assignment with Regression Analysis (1v3)

	Linear regression	Multivariate linear regression
Estimated impact on mothers' total knowledge (Wave=4)	4.446708 (3.742002)	8.363287 (2.355382)
Estimated impact on mothers' total knowledge (Wave=6)	1.391014 (3.848771)	3.776416 (3.001787)
Estimated impact on mothers' total knowledge (Wave=7)	8.334495 (3.219825)	13.03423 (3.22188)

The impact is an increase in score of 4.446708 over the first 6 months and an increase in score of 8.334495 over the third 6 months. Then we run a multivariate regression analysis that controls for some other observable characteristics of the sample groups, and we find that the program has improved the total knowledge of mothers by 8.363287 over the first

6 months and by 13.03423 over the third 6 months, which are different from the linear regression results.

As a result, for the comparison between group1 and group 3, some other factors are systematically different between the treatment and comparison groups that might also explain the difference in mothers' total knowledge. As discussed before, the two groups did not start off with very similar average characteristics. Thus, although they have been exposed to the same set of backgrounds and programs during the several six months of treatment, we cannot say that the treatment is the only plausible reason for the outcome.

### Difference-in-Difference Analysis

The Wide dataset was generated to find the DID (Difference-in-Difference) from baseline to follow-up. Since we divided three kinds of treatment groups and comparison groups (i.e. the comparison between Group1 and Group2, the comparison between Group1 and Group3, and the comparison between Group2 and Group3), we computed the DID for each comparison.

First, we generated variables called DiffCP1v2W1W4, DiffCP1v3W1W4, DiffCP2v2W1W4 to represent the difference of total knowledge from baseline to 6-month. Then, we regressed the difference of total knowledge on the dummy variable treatment1v2 (equals to 0 when the family is in Group2, equals to 1 when the family is in Group1), treatment1v3 (equals to 0 when the family is in Group3, equals to 1 when the family is in Group1), treatment2v3 (equals to 0 when the family is in Group3, equals to 1 when the family is in Group2), respectively. Following table shows the results we found.

Difference-in-Difference from baseline to 6-month (Wave 1 to Wave 4)

Pair	Coefficient=DID	t-statistics	p-value	95% conf. interval
Group1vs Group2	11.76383	2.81	0.006***	3.46698 20.0607
Group1vs Group3	7.423229	1.82	0.071**	-.64493 15.49139
Group2vs Group3	-4.340599	-1.15	0.253	-11.8186 3.13741

In the comparison between group 1 and group 2, the DID equals to 11.76383 and is significant at 95% significance level. That means offering educational books indeed improved first-time mothers' total knowledge on both nutrition and parenting than offering them non-educational books.

In the comparison between group 1 and group 3, the DID equals to 7.423229 at 90% significant level, which means offering educational books improved first-time mothers' total knowledge than offering them no books.

Additionally, we compared the DID between group2 and group3. The large p-value shows that the effect of offering non-educational books was the same as the effect of offering no books.

To conclude, educational books did improve first-time mothers' total knowledge, while the effects of no books or non-educational books were not that obvious.

### Participation Rate and Drop-Outs

`. tab treatment1v2 WaveSort`

treatment1 v2	Wave sorting order (0 = earliest)				Total
	1	4	6	7	
0	66	57	52	47	222
1	62	55	50	44	211
Total	128	112	102	91	433

`. tab treatment1v3 WaveSort`

treatment1 v3	Wave sorting order (0 = earliest)				Total
	1	4	6	7	
0	70	58	57	54	239
1	62	55	50	44	211
Total	132	113	107	98	450

We calculated the participation rates by group at 6,12 and 18 months using the data from the chart above.

	6 months	12 months	18 months
Group1	88.71%	80.65%	70.97%
Group2	86.36%	78.79%	71.21%
Group3	82.86%	81.43%	77.14%

Now we need to understand whether the dropped units are at random – otherwise, it might lead to a biased estimate of the treatment as well as undermine the internal validity of the whole study. The table below shows the ttest result of selected characteristics between the Drop and Participating groups at baseline.

Drop Vs. Non-Drop			
Characteristics	Difference	T-value	Significance
African American	-0.0139232	-0.1789	No
Amer. Ind./AK Natv	N/A	N/A	N/A
Asian	-0.0119714	-0.7431	No
Caucasian/White	-0.0104099	-0.1391	No
Ntv HI/Oth Pac Isl	N/A	N/A	N/A
Spanish/Hispanic	0.0078616	0.1782	No
college_degree	0.0076773	0.131	No
good_health	0.0202993	0.5277	No
paid_job	0.0560797	0.6956	No
live alone	-0.0089722	-0.1502	No
money made last year	0.1521739	0.2631	No
planned pregnancy	-0.0075472	-0.1165	No
Book giveaway program	-0.0032531	-0.1092	No

As none of the t-value is significant (the t-values are really small), we can conclude that the study did not specifically leave any group of people out of the experiment in order to have a “prettier” result. The units who dropped out of the study are completely random. The internal validity of the study still holds.

#### TOT vs ITT

ITT examines the impact on individuals who were assigned to the treatment group regardless of whether they have received the treatment, whereas TOT examines the impact only on participants who actually received the treatment. In this case, as we have full control over who can be enrolled into group 1, group 2 and group 3 and we have been monitoring them in each of the pre-determined time interval, there does not exist any compliance problem. Consequently, the TOT is the same as the ITT.

### **Conclusion and Recommendations**

#### Key Findings from each analysis

##### Pre-Post:

- Mothers in all three groups have their total knowledge increased over the whole evaluation period, but the mean of total knowledge in group2 and group3 are lower than that of treatment group1, inferring that the program could help mothers increase their total knowledge, and the impact will be increasingly significant as program goes on.
- Outcomes within each of the three groups are affected by factors such as ethnicity, level of education, and income, but we could see that mothers who read parental books have their knowledge increased more significantly compared to other two groups. Mothers read random books might not have significant help in total knowledge increase according to the similar coefficient of round in group2 and group3.

##### RA:



- Group2 is a perfect counterfactual for Group1, but Group3 is not.
- Group1 vs Group2: Both group started off with very similar characteristics. The treatment group has a higher total knowledge because of receiving educational books from the intervention.
- Group1 vs Group3: The two groups did not start off with very similar average characteristics. Thus, although they have been exposed to the same set of backgrounds and programs during the several six months of treatment, we cannot say that the treatment is the only plausible reason for the difference in outcome.

DID:

Educational books did improve first-time mothers' total knowledge, while the effects of no books or non-educational books were not that obvious.

### Conclusion

Overall, the evaluation results of three different methodologies all show that the intervention does have a positive impact on improving mothers' total knowledge. However, since the validity of group3 as a perfect counterfactual is undermined due to the difference in characteristics of the mothers in group1 and group3 at baseline, this intervention does not have as strong internal validity.

### Recommendations

This intervention is generally effective in achieving its goal, but the study could be better structured if we have a better similarity level of mother characteristics at baseline, and hence a perfect counterfactual. Besides, the DID design requires a parallel test; however due to the constraint of the data we are unable to perform the sanity check to see whether these three groups fulfils the requirement. Next time for similar study design, we may need to start collecting data 1 year before the implementation if DID is the methodology of choice.

Also, after a more detailed look into the dataset, we discovered that some mothers are not being asked the same amount of questions in the follow-up periods, so AWO1BPC2 could be better as it only takes the corrected percentage based on total questions actually being asked, instead of the total questions the mother ought to answer.