Housing, Health, and Happiness

Group 13: Wei Deng, Meera Duggal

The paper we are going to replicate is the Housing, Health, and Happiness paper from the American Economic Association, authored by Cattaneo, Galiani, et. al. The purpose of this paper is to investigate the impact of a welfare program in Mexico has on the health and happiness of its participants. Specifically, the program replaces dirt floors in households with concrete. The stated objective of the Piso Firme was to improve the living standards, and especially health standards, of vulnerable groups that live in high-density, low-income neighborhoods. The researchers focus on two cities: Lerdo (control) and Torreón (treatment). These two cities were chosen because, although they are in separate states, they occupy the same urban area, which means they are basically identical in their cultural and natural environments. Thus, by picking these two areas the researchers try to limit the bias between the two areas.

The researchers' primary question is if the intervention of the Piso Firme program had effects on the participants' health and happiness. The factors that the researchers used to investigate this impact on residents included rate of parasitic infections, diarrhea, anemia, child cognitive development, respiratory diseases, and self-reported depression/stress levels, among others. The researchers employ a difference-in-difference approach to measuring the causal impact of the program, fitting three linear regression models to predict a number of different outcomes.

The researchers concluded that the Piso Firma program was effective in reducing incidence of intestinal parasites, which improved the cognitive development of children. They also suggest that anti-poverty intervention programs such as Piso Firma can result in major improvements in child health and development.

Data Description

The authors of this paper utilized two main dataframes for their analysis: household dataframe and individual dataframe.

The household dataframe contains information at the household level from both the 2000 Mexican census and the 2005 survey. This data has 78 different variables, each with an indicator of treatment group (0 for control and 1 for treatment). The census data is from 2000, which is before the cement floors were installed. The survey data is from 2005 which is after the treatment. These two data sets combined indicates the different census blocks and their coordinates as well as the characteristics of these census blocks such as # of people, rooms, poverty etc.

The individual dataframe contains information at the individual level for the 2005 survey. This dataset has 89 variables, all of which is from 2005. As opposed to the household data the different variables describe different individual characteristics in the household. This is important because they wanted to gather happiness and stress levels from the moms and cognitive assessment and parasitic infection levels from the children. With this they will be able to conclude if their treatment was effective. This data also includes the coordinates of the household as well as survey data for the mother, father, and child. This describes things such as diarrhea, anemia, and other illnesses. They also include different dummy variables that they are going to include in their regression model. These two datasets, household_df and individual_df, constitute the basis for the analysis of this paper.

The authors of this paper first wanted to ensure that there was a balance between the control and treatment groups. By comparing the means of 22 variables (that summarize demographics structure, housing characteristics, assets, poverty and education of household members) for both the control and treatment group they ensure that the two groups are balanced. They find that there are no significant differences in the two groups and thus say that their samples are well balanced before the welfare program started.

Summary Statistics

Here we recreate the summary statistics table that the researchers use to investigate the balance between the treatment and control households for 27 different independent variables. These values are taken from 2005, after treatment. These variables are grouped into different categories: household demographic, children's characteristics, housing characteristics, hygienic environment, economic variables, and benefits from public social programs. These variables were chosen because they are thought to be unaffected by the treatment.

Table 1: Household demographics

Variable	Observations treatment	Mean treatment	Standard error treatment	Observations control	Mean control	Standard error control	Mean difference
Number of household members	1390	5.312	0.052	1393	5.374	0.049	-0.063
Head of household's age	1390	37.469	0.352	1393	37.120	0.330	0.349
Head of household's years of schooling	1388	6.115	0.092	1391	6.408	0.090	-0.293
Spouse's age	1390	29.643	0.391	1393	28.772	0.392	0.871
Spouse's years of schooling	1233	6.311	0.086	1211	6.479	0.084	-0.168

Table 2: Characteristics of children aged 0-5:

Variable	Observations treatment	Mean treatment	Standard error treatment	Observations control	Mean control	Standard error control	Mean difference
Age	1980	2.649	0.038	2112	2.579	0.037	0.071
Male(=1)	1980	0.491	0.011	2112	0.517	0.011	-0.025
Mother of at least one child in household present (=1)	1980	0.967	0.004	2112	0.964	0.004	0.003
Mother's age (if present)	1898	27.388	0.144	1992	27.465	0.145	-0.077
Mother's years of schooling (if present)	1896	7.029	0.061	1992	6.910	0.061	0.119
Father of at least one child in household present (=1)	1980	0.798	0.009	2112	0.763	0.009	0.035
Father's age (if present)	1512	30.347	0.197	1525	30.632	0.199	-0.286
Father's years of schooling (if present)	1508	6.826	0.081	1519	7.153	0.078	-0.326

Table 3: Housing characteristics:

Variable	Observations treatment	Mean treatment	Standard error treatment	Observations control	Mean control	Standard error control	Mean difference
Number of rooms	1390	2.081	0.029	1393	1.981	0.028	0.100
Water connection (=1)	1390	0.968	0.005	1393	0.977	0.004	-0.009
Water connection inside the house (=1)	1390	0.512	0.013	1393	0.546	0.013	-0.034
Electricity (=1)	1390	0.986	0.003	1393	0.993	0.002	-0.007
Share of rooms with cement floors in 2000	1390	0.329	0.010	1393	0.327	0.010	0.002

Table 4: Hygienic environment:

Variable	Observations treatment	Mean treatment	Standard error treatment	Observations control	Mean control	Standard error control	Mean difference
Household has animals on land(=1)	1390	0.516	0.013	1393	0.480	0.013	0.036
Animals allowed to enter the house (=1)	1390	0.192	0.011	1393	0.190	0.011	0.003
Uses garbage collection service (=1)	1390	0.788	0.011	1393	0.845	0.010	-0.057
Number of times respondent washed hands the day before	1390	3.748	0.040	1393	3.716	0.040	0.032

Table 5: Economic characteristics:

Variable	Observations treatment	Mean treatment	Standard error treatment	Observations control	Mean control	Standard error control	Mean difference
Total household income per capita	1389	1021.654	70.815	1391	1051.676	113.977	-30.022
Total value of household assets per capita	1389	22413.900	200.483	1393	22032.320	196.085	381.580

Table 6: Public Social programs:

Variable	Observations treatment	Mean treatment	Standard error treatment	Observations control	Mean control	Standard error control	Mean difference
Transfers per capita from government programs	1389	16.304	1.128	1392	12.604	0.965	3.699
Household beneficiary of government milk supplement program (=1)	1390	0.059	0.006	1393	0.082	0.007	-0.023
Household beneficiary of government food program (=1)	1390	0.036	0.005	1393	0.022	0.004	0.014

From these tables we can see that household, hygienic environments, socioeconomic status and households that benefit from public programs are balanced between our control and treatment groups. One thing to note is that there is a larger proportion of males in the control group compared to the that of the treatment group. Also, the proportion of children whose father is present in the treatment group is higher than that of the control group. Thus, these variables will be controlled for when we further our analysis.