Socio-economic study on disparities inside a mexican slum

Arthur Valentin, Ilaria Selle, Alice Patin

December 19, 2020

Abstract

Although the program is effective, it's not clear who benefits the most. Are there heterogeneous effects? To answer this question we search correlations between the variables on which the article showed that *Piso Firme* had a positive effect and engineered variables. Looking at the correlations, we find some interesting results concerning the number of children per household and the financial situation of the household. These results are even representative of worldwide trends. Aside from that, still focused on social repercussions, we investigated whether or not *Piso Firme* had a Placebo effect using bootstrapped confidence intervals.

1 Introduction

In these days of pandemics, global health in developing countries is less on the front of the stage. However, global health organizations and countries are still fighting for the well-being, health and mental, of citizens. Housing quality is one their main focus, as it has been shown to impact both health and mental health of children as of the mothers. The analysis done by Matias D. Cattaneo et al. focused on a specific way of improving housing quality in developing countries, such as Mexico: implementation of cement floor in households. Our study aims at completing and extending the analysis of the *Piso Firme* impact, by analyzing whether the effects of the initiative are heterogeneous. Looking at the correlations between engineered features and social variables, we spot a strong correlation between the number of children and the financial situation of a household. This leads us to extend the analysis on the the different factors that can explain this correlation and we compare this result to the worldwide situation, finding meaningful parallels. Going back to the effects of the program, we investigate the possibility that Piso Firme might have a Placebo effect.

2 Data sets

2.1 Household dataset

This is the dataset provided by the paper and it contains all the informations at the household level. It includes data from both the 2000 Mexican Census and the 2005 Survey.

2.2 Children per woman by GDP per capita

This data set is provided by "Our World in Data". It provides us with some insightful informations on the GDP per country in the world and the number of children per woman per country.

2.3 Child mortality GDP per capita

This data set is provided by "Our World in Data". It provides us with some insightful informations on the GDP per country in the world and the child mortality per country These 2 latter datasets serve as comparison to the results found in the first dataset.

3 Methods

We started our analysis by generating some additional variables from the original *Household* dataset. The aim was to derive from the raw variables new socio-economical features, in order to search a correlation with some outputs from Table 6, i.e. stress and depression variables. Therefore, we created new variables such as the parents age, the access to the primary need, the financial health, the proportion of men and children in the household as well as overcrowding, representing the ratio between the number of people in the house and

the number of rooms.

After having computed the correlation between these variables (cf. Figure 1), both with Pearson and Spearman methods, we found almost no statistically significant correlation between the features used. However, we observed that one variable, i.e *number of children*, is highly correlated with the *financial-health*. Moreover, this correlation is statistically significant and negative, meaning that with the decrease of financial health of the household, the number of children increases.

Therefore, going deeper in our research, we focused on the relationship between these two variables. For the analysis, we plotted some others variables, which were also correlated with the *number of children* (but with a weaker correlation coefficient), in order to see if the variable that influenced the most was the *financial-health* or if is the *number of children* which plays a fundamental role in the previous correlation.

On the Figure 2 is shown the relation sheep between these two variables with the highest correlation: the *financial-health* and *number of children*.

We also plotted *1 / financial-health* versus the *number of children* with different slices of parents age and trained linear models for each of this plots.

At this point of the research, we wanted to see if this particular correlation (cf. Figure 2) was specific to that situation (a local population in Mexico) or if it could represent a more general trend. In order to verify this hypothesis, we followed the same manipulation as before with the data, with the only difference that, this time, we used the data extracted from the the other datasets. We replicated the same plot as before, in order to be able to compare a local trend (observed in the previous analysis) with another possible trend (obtained with the data on a global scale).

Finally, we looked into the possibility that *Piso Firme* had also a placebo effect. We computed the beta coefficients associated to the regression y = f(x) with x being the difference in share of cement floor in the household between 2000 and 2005 and y being either the depression or the stress. Then we calculated the confidence intervals at 95% for the difference in betas to evaluate whether or not the placebo was plausible (if the confidence intervals contain 0, then the placebo is very unlikely at a 95% confidence level).

4 Results

4.1 Local results

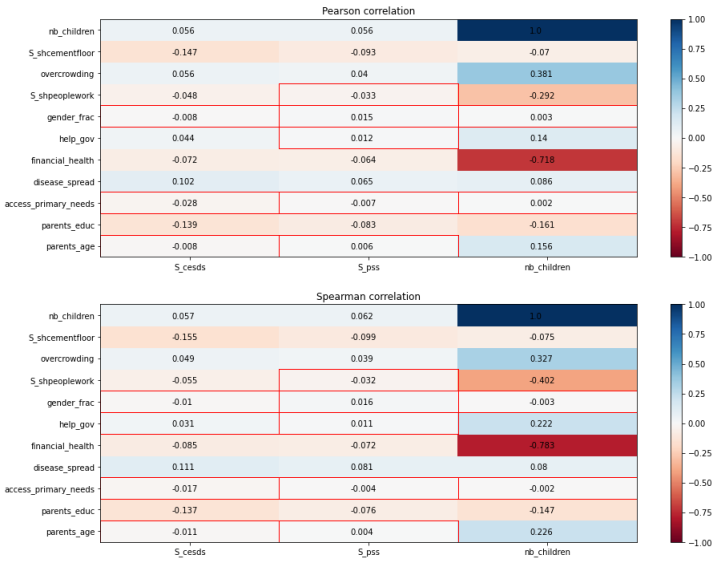

Figure 1: Correlation between the variables, regrouped in different fields

Figure 1: In this figure we note that the only correlation really strong and significant (both with Pearson and Spearman) is between the number of children and the financial-health. Moreover, the relationship between these two variables is negative.

Figure 2: We observe a trend between these two variables, which follow a 1/x function. Moreover, the colors indicate the range of the age of the parents and will help our study for the next comparisons.

Figure 3: We can observe that there exist a linear regression between the number of children and the financial health, and this correlation is maintained in the various slices of age. Moreover, this correlation fits well the data since the R square value is bigger than 0.615 (cf. juppyter Notebook for the exact values).

4.2 World wise results

Figure 4: The trends that we observe at the mexican slum level can be observed at the world level. Here, instead of using the financial health we use the GDP of the country and instead of using the number of children per household, we use the average number of children per woman in the country (or the child mortality, which has the same behaviour as the number of children per woman).

Figure 5: In these regression, we confirm that we observe the same trends previously studied at the local level of a mexican region.

4.3 Placebo effect

To verify our hypothesis of a possible placebo effect done by the treatment (i.e. receive the *Piso Firme* program) on the population, we compute the beta coefficients corresponding to the regression between the difference (from 2000 to 2005) of the *Sshcementfloor* variable and the stress and depression indicators, for both the control and the treatment group. In order to compare these beta we used a bootstrapped confidence interval and we found two intervals for the difference of the beta (one for the prediction of depression and one for the prediction of stress). However, both the confidence interval contain zero, hence we cannot reject that the two beta coefficients are similar at a 95% level. Therefore, we can doubt with a good probability the possible placebo effect of the treatment.

4.4 Conclusion and further studies

First of all, we can conclude that the program had not a placebo effect on the population receiving the treatment.
However, this study brought us to an unexpected interesting finding: we have observed that a socio-economic trend in a population relatively homogeneous (because accurately selected for the proposal of the research carried out by Matias D. Cattaneo et al.) reflects the same trend in a more globalized population (universal scale). Therefore, this means that we could look into ways to transpose measures (e.g. against poverty) from the local to the global scale, and viceversa.

5 References

Cattaneo, M. D., Galiani, S., Gertler, P. J., Martinez, S., Titiunik, R. (2009). *Housing, health,and happiness*. American Economic Journal: Economic Policy,1(1), 75–105.

Our World in Data: https://ourworldindata.org/grapher/children-per-woman-by-gdp-per-capita?tab=chartcountry=region=World

Our World in Data: https://ourworldindata.org/grapher/child-mortality-gdp-per-capita

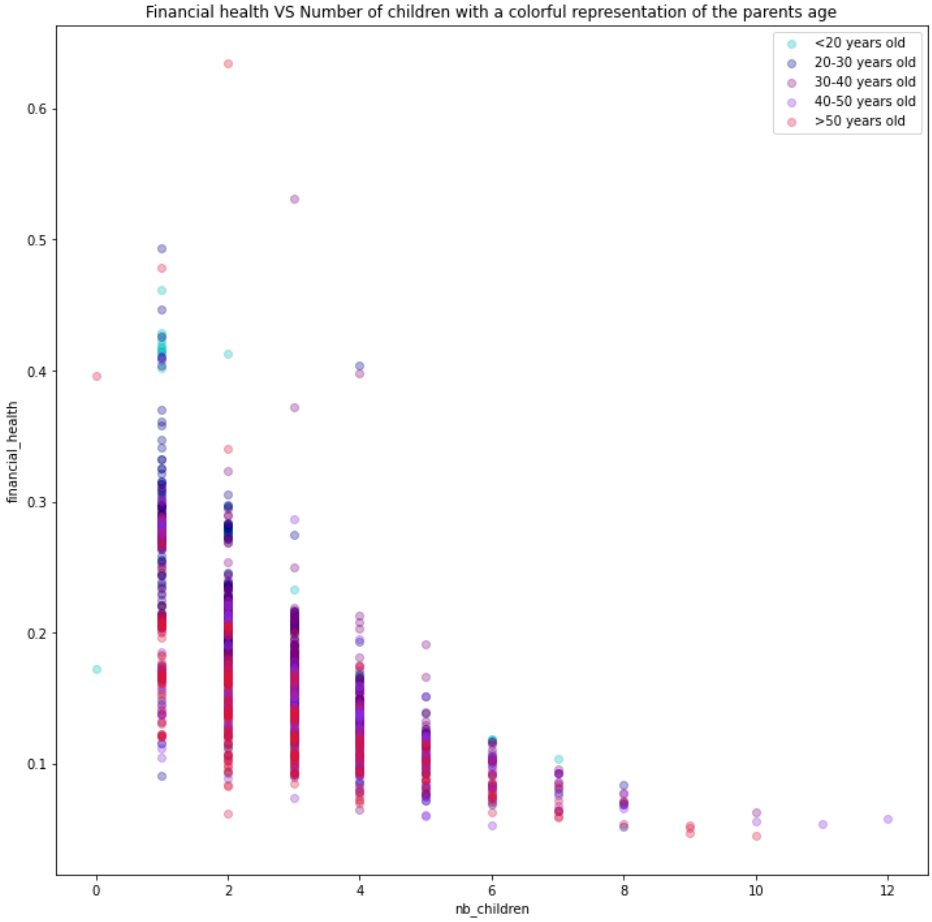

Figure 2: Financial health variable vs. Number of children variable

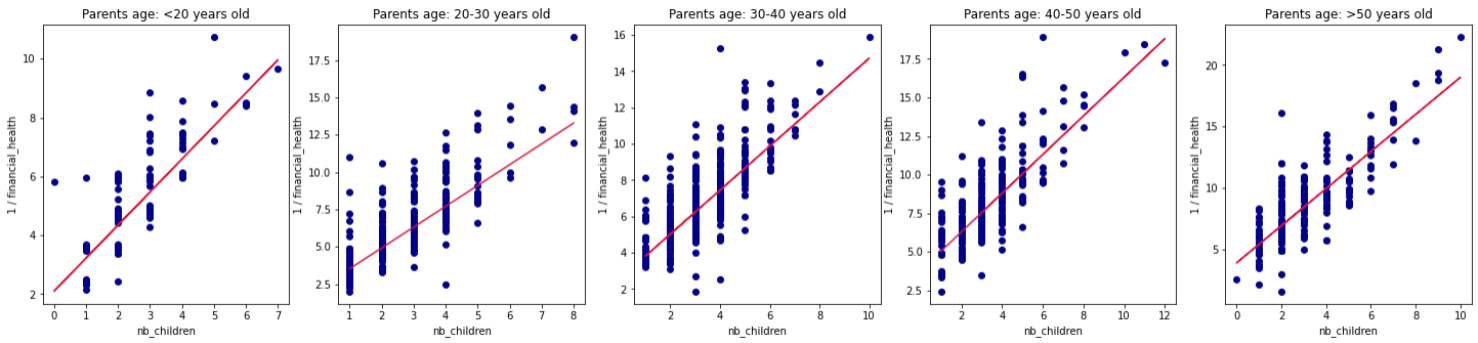

Figure 3: Regression for the features on the Household data level

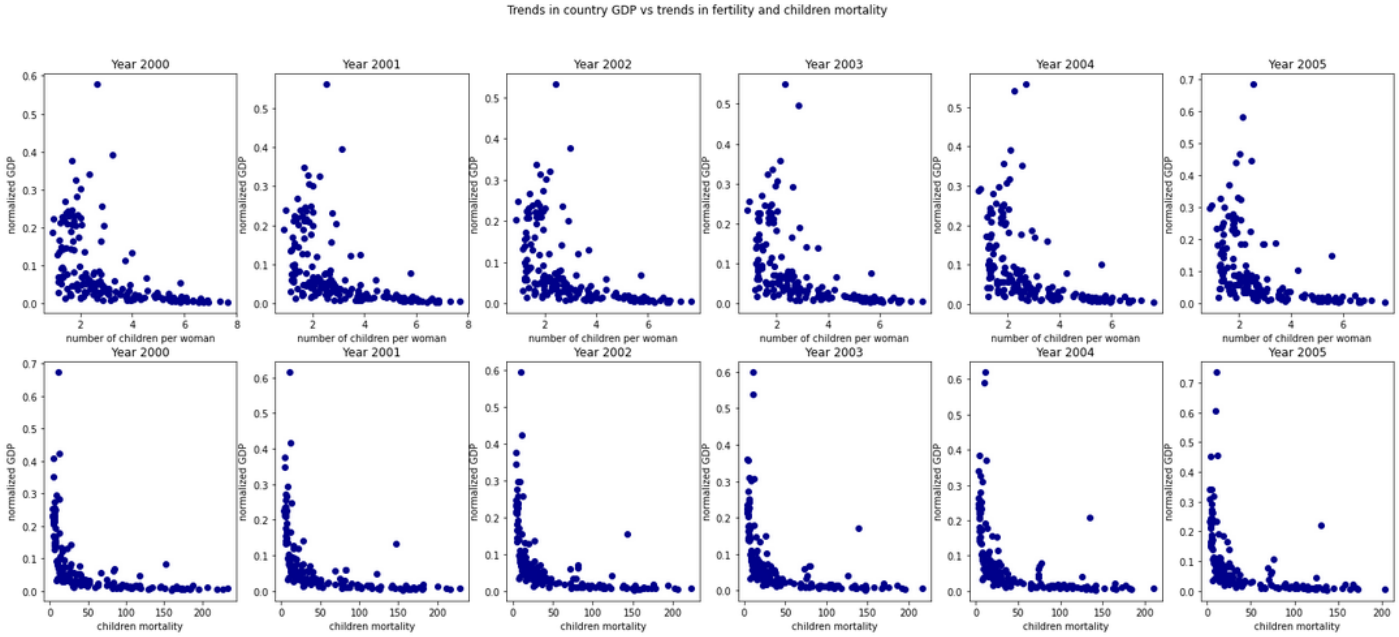

Figure 4: Trends at the world wise level

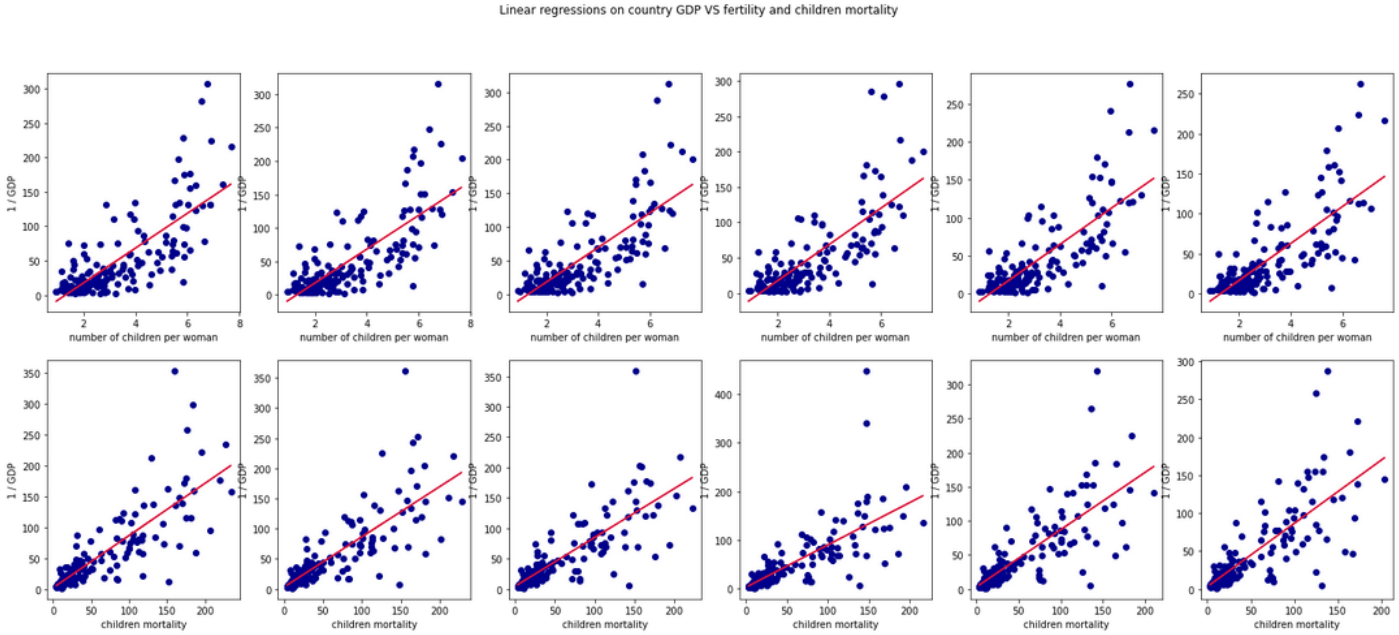

Figure 5: Regression for the same features on the worldwide level