Housing, Health, and Happiness (Replication)

Group 13: Meera Duggal and Wei Deng

May 2022

1 Introduction

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 para- sites, 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.

2 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.

3 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.

4 Statistical Assumptions of Linear Models

A linear model can be described as the following:

$$y_i = \beta_0 + \sum_{k}^{p} \beta_k X_{ik} + \epsilon_i$$

The key statistical assumptions for linear models are:

- Linearity: The relationship between the predictor variables and the response variable is linear.
- Independence: Independence of both your predictors and your errors in the model
- Normality: The residuals terms are normally distributed.
- Homoscedasticity: The variances of the residual is constant for any observation.

4.1 Linearity

Linearity is satisfied when there is a linear relationship between a predictor variable, X, and the response variable y. This assumption can most clearly be checked visually by plotting the predictor variable and the response variable on a scatterplot and noting whether or not the points form a line. If the data is not linear, we can apply an appropriate transformation to satisfy this assumption, like the log transformation.

In the paper, the authors do not explicitly check for the assumption of linearity. The authors do not plot any scatterplots to check for a linear relationship between the dependent variables and the independent variables. That being said, the author does use different factors such as age, demographic, health-habits, and public social programs to explain the given dependent variables.

4.2 Independence

In regards to the independence of the predictors, the variables used in the paper's model are not independent. The authors use variables such as age, demographic, health-habits, and public social program controls. Within these groups they include multiple variables, many of which have a strong association with one another.

Independence of errors can be verified/satisfied primarily during the data collection process. Ideally, the data are collected through the principles of experimental design, where observations for the treatment and control groups are randomly selected. This can be mathematically shown as:

$$\text{cor}(\hat{y}, \hat{\epsilon}) = 0$$

The study followed experimental design principles. The samples for the treated and control groups were randomly drawn from administrative records. The random assignment was predicated by a check of criteria that the authors imposed in order to ensure that households were eligible (for example, the household must have at least one room with a dirt floor). As the basis for their causal inference, the authors ensured that the treatment and control groups were well balanced and investigated the mean difference in variables to check that the differences are not significant between the treatment and control groups.

Additionally, since we are dealing with clusters it's important to discuss the errors within the same cluster and between the different clusters. The paper states that the errors in a single cluster are not independent but they are independent between the different clusters. Thus, this assumption is satisfied.

4.3 Normality

Normality is satisfied when the residuals, $\hat{\epsilon}$, are normally distributed: $\hat{\epsilon} \sim N(0, \sigma^2)$. This assumption can be verified visually by using a Q-Q plot. If the observed and theoretical quantiles fall on a relatively straight line, then the residuals are normally distributed. We can also use formal hypothesis testing to check normality. The Shapiro-Wilk and Jarque-Bera tests are such examples normality testing.

The authors do not explicitly check for normality of the residuals in the model nor do they state that this assumption is being met.

4.4 Homoscedasticity

The assumption of homoscedasticity is satisfied when the residuals of the fitted model, $\hat{\epsilon}$, are constant for every independent variable x. Thus, this is satisfied when

$$\text{cov}(\hat{\epsilon}) = \sigma^2 I_n$$

This assumption can be checked by plotting a scatterplot of the fitted y values and the residuals of a particular x independent variable. If the residuals are randomly distributed across the horizontal line, then the assumption of homoscedasticity is met. However, heteroscedasticity can be seen if the residuals form a pattern on the scatterplot.

The assumption of homoscedasticity for the variance is violated in the paper. Since the standard errors in the paper are clustered at the census-block level, the authors assume homoscedasticity within the clusters, but heteroscedasticity between the clusters. The result is 136 clusters, a large enough number of clusters needed for asymptotic inference. Additionally, the authors suggest that the findings are valid if robust standard errors are computed instead of cluster standard errors.

5 Replication of Main Results

In this section we replicate the main results as they are described in the paper. Model 1 estimates the treatment effect on the response variable with out any control variables. Model 2 now controls for several demographic and health variables. These categories include variables such as: household size, number of rooms, years of education, etc. Model 3 additionally controls for social programs such as household benefits from the government food program, government health insurance program, etc.

The paper accounts for the fact that households may install cement floors on their own, over time. While exploring this possibility they find that households in the control area had a 33% increase of cement floors from the years 2000 to 2005. With that considered, Piso Firme, had positive effects across all domains studied.

Dependent variable	Control group mean (std. dev.)	Model 1	Model 2	Model 3
Share of rooms w cement floors	0.728 (0.397)	0.202 [0.021]*** 27.664	0.207 [0.019]*** 28.503	0.210 [0.019]*** 28.825
Cement floor in kitchen	0.671 (0.538)	0.255 [0.025]*** 37.944	0.260 [0.023]*** 38.734	0.265 [0.023]*** 39.412
Cement floor in dining room	0.709 (0.530)	0.210 [0.025]*** 29.661	0.217 [0.025]*** 30.636	0.221 [0.025]*** 31.204
Cement floor in bathroom	0.803 (0.496)	0.101 [0.022]*** 12.586	0.111 [0.018]*** 13.815	0.114 [0.018]*** 14.169
Cement floor in bedroom	0.668 (0.554)	0.239 [0.020]*** 35.776	0.246 [0.020]*** 36.914	0.246 [0.021]*** 36.900

Missing covariates were imputed with a value of 0. Model 1: no controls; Model 2: age, demographic, and health-habits controlled; Model 3: age, demographic, health-habits, and public social programs controlled. The data given in the table are: estimated coefficient, clustered standard error at census-block level in brackets, and 100 x coefficient/control mean

From the table above, we see that Piso Firme had positive effects on the installation of cement floors in all the dependent variables. Thus, we confirm that the Piso Firme program helped to increase the installation of cement floors for all different types of rooms in a household. We see that across each of the models for each dependent variable, there is a 0.01 level of significance. Also, we see that control and treatment groups are well-balanced.

In the table below we explore how Piso Firme has influenced the health outcomes of young children.

Dependent variable	Control group mean (std. dev.)	Model 1	Model 2	Model 3
Parasite Count	0.333 (1.291)	-0.061 [0.032]** -18.217	-0.064 [0.030]** -19.181	-0.063 [0.030]** -18.999
Diarrhea	0.142 (0.606)	-0.018 [0.009]* -12.445	-0.019 [0.009]** -13.028	-0.017 [0.009]* -12.092
Anemia	0.426 (0.899)	-0.083 [0.028]*** 19.446	-0.074 [0.027]*** -17.476	-0.077 [0.027]*** -18.095
MacArthur Communicative Development Test score	13.354 (94.112)	4.037 [1.631]** 30.230	4.598 [01.582]*** 34.429	4.600 [1.565]*** 34.447
Picture Peabody Vocab Test percentile score	30,656 (72.699)	2.476 [1.679] 8.077	2.715 [1.534]** 8.857	2.615 [1.517]** 8.532
Height-for-age-z-score	-0.605 (1.997)	0.005 [0.042] -0.784	-0.003 [0.040] 0.427	0.001 [0.041] -0.339
Weight-for-height-z-score	0.125 (2.040)	0.012 [0.035] 9.900	0.004 [0.036] 3.200	-0.001 [0.037] -0.087

Missing covariates were imputed with a value of 0. Model 1: no controls; Model 2: age, demographic, and health-habits controlled; Model 3: age, demographic, health-habits, and public social programs controlled. The data given in the table are: estimated coefficient, clustered standard error at census-block level in brackets, and 100 x coefficient/control mean

From the above figures, we see that the offering of Piso Firme significantly affects variables such as parasite count, diarrhea, anemia, MacArthur Communicative Development Test score, and Picture Peabody Vocab Test percentile score. These results suggest that the Piso Firme program positively impacts children less than 6 years old in aspects of infection rate, physical health, and cognitive development. Specifically, with the help of the program, there is a 18.2 percent decrease in parasite counts. Also, there is 12.4 reduction in diarrhea with in the control group. A 20 percentage decrease of anemia is seen in the control group as well. There are many other benefits that can be outlined in the table. One thing to note is that there is not a huge difference seen in the height or weight area.

This table also looks at the cognitive development and if any improvements were seen in that area. This is done because it is expected that the reduction in anemia improved cognitive development.

The final table below shows the impact of the Piso Firme program on maternal satisfaction and mental health, captured by the dependent variables of different satisfaction ratings scales for depression and perceived stress.

Dependent variable	Control group mean (std. dev.)	Model 1	Model 2	Model 3
Satisfaction with floor quality	0.511 (0.668)	0.221 [0.023]*** 43.286	0.225 [0.024]*** 44.104	0.224 [0.025]*** 43.796
Satisfaction with house quality	0.605 (0.670)	0.095 [0.021]*** 15.670	0.089 [0.021]*** 14.700	0.0857 [0.022]*** 14.160
Satisfaction with quality of life	0.601 (0.667)	0.111 [0.022]*** 18.415	0.110 [0.021]*** 18.368	0.111 [0.022]*** 18.431
Depression scale (CES-D scale)	18.532 (12.586)	-2.207 [0.625]*** -11.908	-2.361 [0.573]*** -12.741	-2.342 [0.567]*** -12.637
Perceived stress scale (PSS)	16.514 (9.457)	-1.721 [0.427]*** -10.421	-1.763 [0.397]*** -10.673	-1.753 [0.399]*** -10.617

Missing covariates were imputed with a value of 0. Model 1: no controls; Model 2: age, demographic, and health-habits controlled; Model 3: age, demographic, health-habits, and public social programs controlled. The data given in the table are: estimated coefficient, clustered standard error at census-block level in brackets, and 100 x coefficient/control mean

Across all the dependent variables and the different models, we see that there is significance at the 0.01 level, indicating that these factors have been positively impacted by participation in the Piso Firme program. There is an increase seen with all the satisfaction checks and a decrease seen in both depression and stress amongst parents.

6 Robustness Checks

This section references Table 7 from the paper. The purpose of these robustness checks is to provide evidence that the outcomes from the Piso Firme program were not caused by other confounding variables. The first robustness check was done to ensure that Piso Firme actually caused improvement as opposed to another public program. The second robustness check was done to see whether the children's well-being improved from the lessening of parasitic infection or economic benefit.

The paper already previously addressed the possible bias that could have been introduced by other state specific programs, demographic, health, and socioeconomic characteristics. With the robustness check presented below, the authors are presenting further evidence that removes this possibility that there are confounding variables. Below, we see a series of falsification tests carried out that reinforce the idea that the Piso Firme program had a significant impact on these different variables.

Table 7

Dependent variable	Control group mean (std. dev.)	Model 1	Model 2	Model 3
Respiratory diseases	0.355 (0.858)	0.0207 [0.0187] 5.812	0.0196 [0.0183] 3.386	0.0179 [0.0189] 3.124
Skin diseases	0.1011 (0.538)	0.0000812 [0.0116] 0.08	-0.001 [0.0111] -0.83	-0.001 [0.011] -1.17
Other diseases	0.0408 (0.363)	0.00538 [0.009] 13.12	0.00526 [0.0089] 10.17	0.005 [0.0089] 9.66
Installation of cement floor	0.53 (0.578)	0.3763 [0.0281]*** 70.98	0.375 [0.0283]*** 90.17	0.376 [0.0281]*** 90.36
Construction of sanitation facilities	0.014 (0.411)	0.0173 [0.015] -17.02	-0.0176 [0.015] 10.18	-0.0168 [0.0149] 10.88
Restoration of sanitation facilities	0.044 (0.29)	0.0014 [0.0126] -3.156	-0.0008 [0.0126] 1.03	-0.0015 [0.0123] 1.894
Construction of ceiling	0.159 (0.535)	0.0284 [0.0239] 17.821	0.0208 [0.024] -20.76	0.0178 [0.0232] -20.26
Restoration of walls	0.111 (0.454)	0.0124 [0.0167] 11.199	0.012 [0.0155] -10.33	0.0144 [0.0156] -16
Any house expansion (excluding installation of cement floors)	0.277 (0.646)	0.0452 [0.0312] 16.313	0.0374 [0.031] -10.7	0.0376 [0.0301] -16.37
Log of self-reported rental value of house	5.918 (1.13)	0.0353 [0.0398] 0.596	0.0528 [0.0311] 1.14	0.056 [0.0314] 1.21
Log of self-reported sale value of house	10.491 (1.761)	-0.0433 [0.099] -0.413	-0.017 [0.083] -0.3	-0.0144 [0.0794] -0.253

Log total income of mothers of children 0-5 years	7.791 (2.86	-0.416 [0.0637] -0.534	-0.039 [0.0634] -0508	-0.035 [0.068] -0.457
Log total income of fathers of children 0-5 years	8.121 (1.543)	-0.0152 [0.0271] -0.187	-0.005 [0.0261] -0.07	0.0006 [0.0255] 0.00839
Total consumption per capita	753.733 (1628.084)	4.27 [43.86] 0.567	12.992 [43.7] -0.0000007	16.15 [42.63] -0.0000009

Missing covariates were imputed with a value of 0. Model 1: no controls; Model 2: age, demographic, and health-habits controlled; Model 3: age, demographic, health-habits, and public social programs controlled. The data given in the table are: estimated coefficient, clustered standard error at census-block level in brackets, and 100 x coefficient/control mean

The table tries to see if there is a relationship between health illnesses that would not be linked to the removal of dirt floors in a house. It is also checking to see if there is a difference between illnesses in the control and the treatment group. In the table above we see that the treatment variables are not highly associated with any of the dependent variables, so we have shown that there are no other out side factors that could impact the results of the Piso Firme study. This robustness check also removes the ideas that treatment and control households value their houses at a higher level and that the houses in both groups were not improved more than the other. It was also found that none of the coefficients in the table were statistically significant. This shows that the bias between the two groups and the impact of possible confounding variables are minimal.

7 Reanalysis

For our re-analysis of the paper we decided to employ leave-one-out estimation and compare our results to the one that are described in the paper. Before, we just fitted the model with all of the observations. For this reanalysis we are fitting the model without a single cluster of observations. We find in this paper that there are a total of 136 clusters, thus the LOOE models are fit with 135 clusters. With these results we would then simply compare the coefficients that are outputted from two different models. As an extension of Theorem 11.2 (which can be further proven for clusters) we know that we only need to refit OLS once by removing only one random cluster from the data.

We are only concerned with the coefficient associated with the `dpisofirme` `dpisofirme` variable, so we have created a comparison between the coefficients for the model without one cluster and the model that has all clusters. We have recreated models 1, 2, and 3 with the same dependent variables as the ones that were listed in the main results section.

Model 1, 2, and 3 have the same control variables as described in the main results section.

- Model 1: Estimates the treatment effect on the response variables without any control variables
- Model 2: Estimates the treatment effect on the response variable while controlling for demographic and health variables

- Model 3: Estimates the treatment effect on the response variable while controlling for different demographic, health, and benefits received from social programs

Dependent variable	Model 1 Coef. Est. (Full Model)	Model 1 Coef. Est. (LOOE)	Model 2 Coef. Est. (Full Model)	Model 2 Coef. Est. (LOOE)	Model 3 Coef. Est. (Full Model)	Model 3 Coef. Est. (LOOE)
Share of rooms w cement floors	0.202	0.202	0.207	0.207	0.210	0.209
Cement floor in kitchen	0.255	0.255	0.260	0.259	0.265	0.264
Cement floor in dining room	0.210	0.209	0.217	0.215	0.221	0.219
Cement floor in bathroom	0.101	0.105	0.111	0.112	0.114	0.116
Cement floor in bedroom	0.239	0.238	0.246	0.245	0.246	0.244

Missing covariates were imputed with a value of 0. Model 1: no controls; Model 2: age, demographic, and health-habits controlled; Model 3: age, demographic, health-habits, and public social programs controlled. The data given in the table are: estimated coefficient, clustered standard error at census-block level in brackets, and 100 x coefficient/control mean

Dependent variable	Model 1 Coef. Est. (Full Model)	Model 1 Coef. Est. (LOOE)	Model 2 Coef. Est. (Full Model)	Model 2 Coef. Est. (LOOE)	Model 3 Coef. Est. (Full Model)	Model 3 Coef. Est. (LOOE)
Parasite Count	-0.061	-0.663	-0.064	-0.068	-0.063	-0.067
Diarrhea	-0.018	-0.018	-0.019	-0.019	-0.017	-0.017
Anemia	-0.083	-0.086	-0.074	-0.078	-0.077	-0.081
MC Dev. Test score	4.037	3.937	4.598	4.704	4.600	4.533
PPV Test percentile score	2.476	2.741	2.715	2.920	2.615	2.823
Height-for-age-z-score	0.005	0.007	-0.003	-0.002	0.001	0.001
Weight-for-height-z-score	0.012	0.003	0.004	-0.005	-0.001	-0.010

Missing covariates were imputed with a value of 0. Model 1: no controls; Model 2: age, demographic, and health-habits controlled; Model 3: age, demographic, health-habits, and public social programs controlled. The data given in the table are: estimated coefficient, clustered standard error at census-block level in brackets, and 100 x coefficient/control mean

Dependent variable	Model 1 Coef. Est. (Full Model)	Model 1 Coef. Est. (LOOE)	Model 2 Coef. Est. (Full Model)	Model 2 Coef. Est. (LOOE)	Model 3 Coef. Est. (Full Model)	Model 3 Coef. Est. (LOOE)
Satisfaction with floor quality	0.221	0.219	0.225	0.223	0.224	0.221
Satisfaction with house quality	0.095	0.091	0.089	0.086	0.0857	0.083
Satisfaction with quality of life	0.111	0.112	0.110	0.111	0.111	0.112
Depression scale (CES-D scale)	-2.207	-2.327	-2.361	-2.430	-2.342	-2.396
Perceived stress scale (PSS)	-1.721	-1.754	-1.763	-1.775	-1.753	-1.758

Missing covariates were imputed with a value of 0. Model 1: no controls; Model 2: age, demographic, and health-habits controlled; Model 3: age, demographic, health-habits, and public social programs controlled. The data given in the table are: estimated coefficient, clustered standard error at census-block level in brackets, and 100 x coefficient/control mean

In the above tables LOOE stands for leave-one-out. By comparing the coefficients from the model with the full data and the model that has a single cluster left out we see that the coefficients are very similar. Thus, this means that a good model was fit in this paper as it can be generalized to all of the clusters (i.e. each of the census blocks).