# Online Appendix

## Relative Price of Sugar and Sugar Production

[[Figure A1 here]]

Figure A1 shows the log-relative price of sugar to coffee from 1890 to 1940. The prices come from the NBER Macrohistory database, Ch IV: Prices. The prices used are the U.S. Wholesale Price of Granulated Sugar (Bureau of Labor Statistics 1/1890-12/1941) and U.S. Wholesale Price of Coffee (New York 01/1890-12/1940). From the graph it is easy to see that from the early 1890s to the early 1900s there was a large increase in the price of sugar relative to the price of coffee.

Figure A2 shows the history of Haitian sugar production as reported in Deerr (1949). From 1710 to 1836, the line is the total tons in both clayed and muscavado production, whereas 1922 to 1943 is raw sugar. The data shows the rise, fall, and eventual return of sugar. Note that in 1789, the numbers reported by Deerr are about half as large as the numbers from Mackenzie (1830), so the Deerr figures might underestimate the relative swings in Haitian production.

## Image Processing Algorithm

[[Figures A2-4 here]]

The image processing algorithm takes three steps, represented in Figure A3. First, the algorithm extracts the simulated plantation from the map created by the 1956 U.S. Army Inter-American Geodetic Survey. Next, the algorithm converts the colored image to black and white, which erases many of the irrelevant features such as contour lines or shading. Finally, the algorithm tags all groups of connected black pixels. The algorithm locates a black pixel, then checks each neighbor to the north, east, south, and west to see if it is also black. Adjacent black pixels are labeled connected, and a group is a set of pixels such that every pixel is connected to at least one other pixel in the set. The algorithm identifies all groups, then it classifies a group as a house if it meets two criteria: (1) the group contains between 8 and 24 pixels inclusive and (2) the farthest distance between any two pixels in the group is less than or equal to 5 pixels. The first criterion comes from the observation that the house markers were typically 3x3 to 5x5 pixels. The second eliminates groups that fit the first criterion but are too diffuse to be a house marker; for example, a road that is 20 pixels long.

The algorithm does not perfectly identify dots in every image. Comparing Step 1 and Step 3 in Figure A1, we can see in the north-east corner of the plantation the algorithm missed four dots where the pixels touched the latitude line. On the other hand, in the same region of the image, the algorithm counted the intersection of the same line with the river as a dot. To check the algorithm’s accuracy, 1,629 images were selected for both hand-counting and the algorithm. Figure A4 shows the relationship between the hand-counted and machine-counted tallies, and the correlation coefficient between the two counts is 0.96.

## Land Rental Data - Sources

[[Table A1 about here]]

Table 6.3 lists the issues of *Le Moniteur* from which the land rental data were collected.

## Lower bound estimate

[[Table A2 about here]]

While I was unable to find information on the supply of land the government held in each district, there was one source that can give us a lower bound. In 1934, the government listed properties available for rent in five districts (*Le Moniteur*, 1934 No. 24, 22 March 1934). Across the 22 years of gazettes, this was the only example of the government promoting land, but it shows us at least what was available in one year. Table A2 shows that in four of the five districts farmers did not even adopt as much land in 16 years as was available in the one year.

## Testing Transaction Costs with Microdata

[[Table A3 about here]]

A nice feature of the microdata is that we can test the validity of the transaction cost proxy using the data on neighbors. More houses in the district should increase the probability of finding a private neighbor, and a greater spread should increase the probability too. Table A3 reports the results of regressions with the settlement data using two definitions of a private neighbor. One definition classifies the plot as having a private neighbor if at least one neighbor is not explicitly designated as government land; the second definition looks at all plots with at least one neighbor that is government land and one that is not. When we look at just the relationship between $μ $and private neighbors, we get a positive and statistically significant result, as we expect. When $σ$ is added as a control, the statistical significance disappears. Furthermore, though not statistically significant, the coefficient on $σ$ is positive, which runs counter to the prediction. The regressions with just *µ* are promising, but the regressions including $σ$ are underpowered, so it is difficult to draw strong conclusions from these tests.

Table A1: Year and issue numbers of *Le Moniteur* that provided data on rental plots

|  |  |
| --- | --- |
| Year | Issue Numbers |
| 1929 | 24; 55; 82 |
| 1930 | 55; 105 |
| 1931 | 3; 38; 64; 84; 105 |
| 1932 | 30; 59; 88; 104 |
| 1933 | 27; 62; 64; 100 |
| 1934 | N/A |
| 1935 | 2; 18; 38; 73; 102 |
| 1936 | 19; 56; 84 |
| 1937 | 14; 45; 73; 102 |
| 1938 | 9; 51; 68; 90; 104 |
| 1939 | 9; 30; 50; 81 |
| 1940 | 10; 50; 76; 79 |
| 1941 | 3; 41; 70; 94; 108 |
| 1942 | 10; 30; 40; 70; 100 |
| 1943 | 26; 60; 85; 97 |
| 1944 | 12; 53; 72; 90; 95; 103 |
| 1945 | 28; 29; 41; 49; 55; 72; 80; 84; 106 |
| 1946 | 13; 17; 46; 62; 64; 65; 99; 124; 125 |
| 1947 | 10; 11; 12; 56; 63; 103; 115 |
| 1948 | 23; 43; 100 |
| 1949 | 23; 43; 71; 103 |
| 1950 | 1; 7; 32; 79; 83; 106; 148 |

Notes: The 1934 Moniteur was presented in a single volume without issue numbers.

Table A2: Comparing Land Available in 1934 to All Land Rented from 1934 to 1950

|  |  |  |  |
| --- | --- | --- | --- |
| District | Department | Available | Rented |
|  |  | (1934) | (1934-1950) |
| Croix-des-Bouquets | Ouest | 470 | 94 |
| Fort Liberte | Nord | 1,593 | 1,077 |
| Thomazeau | Ouest | 223 | 610 |
| Petionville | Ouest | 301 | 37 |
| Ganthier | Ouest | 168 | 532 |

Notes: All figures are in hectares. Available land comes from a 1934 advertisement published in *Le Moniteur*, 1934 No. 24 (22 March 1934).

Table A3: Using microdata to test hypotheses about transaction costs

|  |  |  |
| --- | --- | --- |
|  | *P* (*Private N eighbor*) | *P* (*Private Neighbor|One State*) |
| ln(*µ*) | 0.19\*\*\* | 0.11 | 0.16\*\* | 0.11 |
|  | [0.06] | [0.13] | [0.06] | [0.14] |
| ln(*σ*) |  | 0.09 |  | 0.06 |
|  |  | [0.13] |  | [0.15] |
| Observations | 4,805 | 4,805 | 2,927 | 2,927 |
| R-squared | 0.05 | 0.05 | 0.04 | 0.04 |

Notes: Standard errors clustered by district.

Figure A1: The Price of Sugar Relative to Coffee, 1890-1940



Notes: Data come from NBER Macrohistory: IV Prices.

Figure A2: Haiti sugar production, 1710-1943



Sources: Data come from Deerr (1949).

Figure A3: Example of a simulated plantation going through the image processing algorithm



Notes: Step 1 shows a simulated plantation in the Artibonite department extracted from 1956 U.S. Army Inter-American Geodetic Survey. The simulated plantation is 256x256 pixels, or roughly 100 hectares. Each dot on the plantation represents a building, usually huts. In Step 2, the algorithm converts the image to black and white to remove extraneous details, such as the water. In Step 3, the algorithm finds all groups of black pixels that match the criteria for a house, as described in the text. The algorithm counts all groups black pixels in the Step 3 image and concludes there are 45 buildings on this simulated plantation.

Sources: United States Army Map Service (1962).

Figure A4: Comparing machine-counted dots to hand-counted dots



Notes: Each dot represents a simulated plantation (N=1,629). The x-axis represents the number of dots that the image processing algorithm counted on the simulated plantation. The y-axis represents the number of households that a research assistant counted. The Pearson correlation coefficient between the two counts is 0.96.

Sources: See text.