Replication Readme for

"The hidden cost of bananas: The effects of pesticides on newborns' health"

Conditionally Accepted at the the Journal of the Association of Environmental and Resource
Economists
Joan Calzada (calzada@ub.edu)
Meritxell Gisbert (meritxell.gisbert@uab.cat)
Bernard Moscoso (hmoscoso@espol.edu.ec)

This file contains information on the programs and data used to generate the tables and figures
included in the published paper and in the Online Appendix. The specific techniques used to collect
and construct the data are described in the published paper and in Section C of the Online Appendix.
The following files have been uploaded to the Harvard Dataverse archive:

Files:

"Readme_Banana_Data_Replication.pdf" (this Readme document)
"Bananas.do" Stata Do file to generate figures and tables
"bananas1.dta" Stata data file
"bananas2.dta" Stata data file
"bananas3.dta" Stata data file

Data files:

"bananas1.dta" Stata data file that contains the final dataset used to generate most of the figures
and tables included in the published paper and in the online appendix.

"bananas2.dta" Stata data file that contains the dataset used to generate the tables with
maternal fixed effects included in the published paper and in the online appendix.

"bananas3.dta" Stata data file that contains the dataset used to generate the tables in the online
appendix that examine the effects of pesticides in other crop plantations.

The three datasets contain all the variables that reflect the mother's and newborn's
characteristics used in our estimations. Datasets "bananas1" and "bananas2" contain the two
main variables used in our causal analysis of Section 5: the dummy variable *Banana Exposure*
determines the group of newborns geographically exposed to pesticides; and the dummy
variable *Intensive Fumigations* reflects the newborns gestated in periods with intensive

fumigations. The variable *Banana Exposure* identifies 8,865 newborns geographically exposed to fumigations, and 41,181 newborns geographically non-exposed, within 2,500 meters of the closest banana plantation. Notice that our empirical analysis focusses on 151 grid cells that exhibit variability in the dosage of pesticides above and below the 4 gallons per hectare per month threshold. As a result, our final sample includes 5,707 geographically exposed and 15,686 non-exposed newborns. The dataset "bananas3" contains a broader subsample of newborns which are considered in the estimations of Section 6 that examine the robustness of our results. We use data from the aerial fumigation registry showing whether newborns were directly exposed to fumigations. Moreover, we test the robustness of our results to different periods and intensities in the exposure to pesticides, and we apply a placebo test that considers newborns who were geographically exposed to other crops (rice, cocoa, and corn) and seasonally exposed to banana fumigations. This dataset considers a sample of 136,070 newborns overall, from which different subsamples are used for different tables.

Data sources:

The Stata data files contain variables from the sources we explain below. These datasets are in the public domain and can be distributed with proper citation:

1. Birth data

Birth data for the period 2015-2017 was obtained from the INEC:
https://www.ecuadorencifras.gob.ec/nacimientos-y-defunciones-informacion-historica/
This dataset is publicly available, but it does not contain the newborns' address during pregnancy. To obtain this additional information it is necessary to request it to the INEC, through their official email address: sugerencias@inec.gob.ec or to inec@inec.gob.ec. The collection of this confidential data and its manipulation is only possible in the laboratory of the INEC, located in its headquarters in Quito (Ecuador).

We used this dataset to create the variables about the mothers and the newborns characteristics. Using Geographic Information System tools (GIS) with the software QGIS, we georeferenced the location of newborns using the mother's address during pregnancy. This process implied several APIs from Google maps, governmental source maps, and open street maps. Afterwards, we computed the grid cell in which newborns were georeferenced and we generated the grid cell identifier.

Finally, we merged this dataset with the datasets on plantations and pesticides described below. This allowed us to create the variables that describe whether newborns were geographically exposed to the banana plantations and whether they were gestated in periods of intensive fumigations. The exact procedure followed in QGIS is detailed in Section C in the Online Appendix.

2. Banana plantations data

Banana plantations data was obtained from the *Ministerio de Agricultura y Ganadería* (http://geoportal.agricultura.gob.ec/). Data for all crop plantations was obtained from the Open Data Catalog: https://datosabiertos.gob.ec/dataset/mapa-de-catastro-bananero-en-el-ecuador-continental-escala-1-5-000-ano-2013

The banana plantation dataset uses information collected in a census that took place in 2013, which provides information on the type of fumigation applied to each plantation (i.e. manual or aerial fumigations). More importantly, it contains the coordinates of the polygon of the plantation. The dataset can be projected using the Coordinate Reference System of 17S – Ecuador.

We used this dataset to compute the distance from the newborns' residence to the plantations. Moreover, we constructed 25-meter-radius buffers from the newborns' residence up to 2.5 kilometers, and for each buffer we calculated the area in square meters of aerially fumigated plantations. These calculations were performed with QGIS. For a detailed explanation see Section C in the Online Appendix.

Data on all plantations consists of a satellite map of every crop in Ecuador, and it contains georeferenced information of the perimeters of the plantations. The process described above was used to compute the distance from the newborns' residence to the plantations and the buffers with the square meters of other crops plantations.

3. Pesticides data

Information about the fumigations of banana plantations has been obtained from the *Ministerio de Medio Ambiente, Agua y Transicion Ecologica*. The generic link to access this information is: http://pras.ambiente.gob.ec/web/siesap/agricola-y-pecuario

Georeferenced data about the use of pesticides in the plantations can obtained upon request to the responsible persons of the "*Programa de Reparación Ambiental y Social PRAS*", through their official email address: pras@ambiente.gob.ec. More information can be found in the following webpage: https://www.ambiente.gob.ec/programa-de-reparacion-ambiental-y-social-pras/

We used QGIS to merge this dataset with the plantations and birth datasets. Afterwards, we used STATA to compute the pesticides profiles at the grid cell and month level and we assigned the profiles to the newborns using the grid cell identifier.

The pesticides profiles show the average gallons of pesticides per hectare applied in each grid cell per month. They are defined as the sum of the gallons of pesticides sprayed in the grid cells

per month, divided by the number of hectares with banana plantations. Section B in the Online Appendix B shows the evolution over time of pesticide use in the grid cells examined in our paper. More information on the process for defining the periods of intensive fumigations at the grid cell and month level can be found in the published paper.

4. Other crop plantations

Information about the other crops considered in the paper can be found in the following link of the *Ministerio de Agricultura y Ganadería*:
https://datosabiertos.gob.ec/dataset/mapa-de-estimacion-de-superficie-sembrada-de-los-cultivos-anuales-primer-periodo-siembra-ano-2021

We used this dataset to compute the distance from the newborns' residence to the other plantations. Moreover, we created buffers from the newborns' addresses that reflect the number of square meters of other crop plantations that were close to the newborns.

For further inquiry, revision, and clarification on data collection, please contact the authors.