Testing the effects of mobile health in Ecuador

Alonso Quijano

5/15/2022

Ecuador expanded its primary healthcare access program using mobile health teams through the Programa Médico de Barrio (PMB), which was rolled out gradually across cantons. A team composed of a physician, a nurse and a data specialist made house visits to marginalized and poor families in rural and peri-urban areas. Using two nationally-representative household surveys, this study estimates the impact of the PMB on various outcomes: reported health problem (in the 30 days prior to the survey), healthcare use to address the health problem, hospitalization, preventative care and self-reported health status. The independent variable of interest is any exposure to PMB. Alternatively, we also use as exposure the proportion of potential weeks covered by PMB to compare outcomes of treated vs. non- or partially-treated cantons based on the random combination of timing of program implementation and timing of survey interview.

Data

Data were derived from two nationally-representative health and nutrition surveys, Encuesta de Salud y Nutrición (ENSANUT), which were fielded during 2012-13 and 2018-19. The first survey was collected from 92,500 individuals under the age of 60; topics included anthropometric, blood and urine measurements, tobacco and alcohol use, physical activity, diet (through a 24-hour food recall diary), as well as issues of health care use and access (https://ghdx.healthdata.org/record/ecuador-national-health-and-nutrition-survey-2012). The second survey collected data from 168,747 people during November and December 2018 (for 82% of the sample) and during January to July 2019 (for remaining 18% of the sample). The ENSANUT surveys are publicly available at: https://www.ecuadorencifras.gob.ec/salud-salud-reproductiva-y-nutricion/.

The mobile health teams program (PMB)

In August 2017, President of Ecuador Lenin Moreno launched the "Programa Médico de Barrio" (hereafter, PMB) [neighborhood doctor program] to strengthen primary care with a focus on health promotion and disease prevention, as a tool for citizens to access health services (Ministry of Health, 2019). The plan would progressively expand the program to cover the entire country with mobile teams. Each team was composed of three members: a general internal medicine physician, a person with a bachelor's degree in nursing, and a technician in primary care. The stated PMB's strategic objectives included: (a) to re-orient a curative approach, centered on disease and the individual, towards producing better health based on primary care; and (b) to bring health services closer to the community, thereby reducing access barriers (Ministerio de Salud Pública, 2017).

Outcome varaibles:

The main outcomes were reported health problem in the 30 days prior to the survey date (*hproblem*); and healthcare utilization to address the health problem (*gotcare*).

Specifically, *hproblem* asked:
Did [person j] in the last 30 days [from ... to...] have any illness, accident, burn, toothache, earache or any other discomfort, even if it was temporary?

And *gotcare* asked:
What did [person j] do as the first action to solve the (health problem)? 1. Visited a hospital, dispensary, health center or subcenter; consulted a doctor, healer, etc.? 2. Got care at home from a doctor, nurse, healer, etc.? 3. Self-medicated? 4.Had to be admitted to a hospital, clinic, etc.? 5.Nothing. *Gotcare* was recoded as dichotomous variable with 0=nothing or self-medicated, and 1 if otherwise. Note that for those who did not have a health problem (i.e., if *hproblem*=0) the variable *gotcare* was missing.

For preventative care the question was: Now I am going to ask you questions about preventive care, in the last 30 days (from.....to...), were you checked by a psychologist, dentist, healer, apothecary or massage therapist, did the neighborhood doctor visit you at home, or did you receive any preventive service such as: vaccinations, well-child check-ups, blood pressure checks, dental check-ups, etc.? And for health perception (hpercep) the original question was: How would you rate the current state of health of (...): Excellent?...1 Very good?...2 Good?...3 Fair?...4 Poor?...5

We reversed coded *hpercep* to make interpretation easier; i.e., that higher number implied better health. And we also created a new health perception dummy variable (*hpercepd*) =1 if respondents chose excellent or very good health, and =0 otherwise.

The exposure variable was operationalized in two ways. First, a dummy variable (*pmb*=1) indicated if a canton had at any point been exposed to PMB (prior to the latest ENSANUT survey date of July 2019). The second approach measured treatment intensity via a constructed variable called proportion of potential weeks covered (*pweekscov*). The variable *pweekscov* had as a denominator the maximum potential number of weeks that a canton could have been exposed to PMB (counting from the first PMB rollout date to the latest ENSANUT survey data collection date), and the numerator was the actual number of weeks that the canton was exposed to PMB. Dates when PMB started in each canton were collected from dozens of publically available non-governmental sources (including media outlets, Internet newspapers and social media posts) as well as official governmental posts.

DID model

This paper uses a panel, fixed-effects approach (Angrist & Pischke, 2009, 2015). The unit of analysis is the *cantón* (or county), which is the second highest level of administrative division in Ecuador, after *parroquia* (or parish). There are 221 cantons distributed among 24 provinces. The rollout stages across different cantons leads to a natural implementation of a staggered-entry differences-in-differences (Athey & Imbens, 2022; Callaway & Sant'Anna, 2021) model to evaluate the effects of PMB. Once a canton started the PMB, it was assumed that it was afterwards always exposed. Because of the different dates of PMB start at the canton level and the different dates of data collection for the ENSANUT 2018-19 survey, there is a natural experiment in terms of staggered entry into treatment. The first DID specification is as follows:

$$Y_{it} = \alpha_0 + \alpha_1 PMB + \alpha_2 POST + \alpha_3 PMB \times POST + \epsilon_{it}$$

where Y_{it} indicates the outcome of canton i at time t, PMB_{it} is a dummy variable indicating any presence of the Programa Médico Barrio at the canton level; $POST$ is an indicator (=1) if the time is 2018-19; and where the coefficient of interest is the interaction α_3.

The second DID specification was as follows:

$$Y_{it} = \alpha_0 + \beta_1 pweekscov_{it} + \beta_2 POST + \beta_3 pweekscov_{it} \times POST + \epsilon_{it}$$

where the treatment intensity was measured by the proportion of potential weeks that a canton was covered by the PMB $pweekscov_{it}$ at the time of the ENSANUT 2018-19 survey visit took place, and where the coefficient β_3.

Results

The summary statistics are:

```
# summary statistics
pmb_data %>% select(cantcode2, esdate, pmb_start_date, hproblem, gotcare,
                    hosp, hpercepd, pmb, pweekscov, post) %>% summary()
```

```
##   cantcode2             esdate       pmb_start_date     hproblem
##  Length:342         Min.   :17851   Min.   :17391   Min.   :0.007641
##  Class :character   1st Qu.:17859   1st Qu.:17562   1st Qu.:0.215862
##  Mode  :character   Median :17870   Median :17809   Median :0.313763
##                     Mean   :17891   Mean   :17783   Mean   :0.330968
##                     3rd Qu.:17883   3rd Qu.:17970   3rd Qu.:0.432566
##                     Max.   :18090   Max.   :18248   Max.   :0.723404
##     gotcare            hosp            hpercepd          pmb
##  Min.   :0.2500   Min.   :0.00000   Min.   :0.00564   Mode :logical
##  1st Qu.:0.5158   1st Qu.:0.02878   1st Qu.:0.13065   FALSE:134
##  Median :0.6334   Median :0.03946   Median :0.18379   TRUE :208
##  Mean   :0.6316   Mean   :0.03941   Mean   :0.19209
##  3rd Qu.:0.7231   3rd Qu.:0.04876   3rd Qu.:0.23966
##  Max.   :1.0000   Max.   :0.12228   Max.   :0.73243
##    pweekscov         post
##  Min.   :0.0000   Mode :logical
##  1st Qu.:0.0000   FALSE:171
##  Median :0.1944   TRUE :171
##  Mean   :0.2447
##  3rd Qu.:0.4620
##  Max.   :0.8158
```

First DID specification model results

$$Y_{it} = \alpha_0 + \alpha_1 PMB + \alpha_2 POST + \alpha_3 PMB \times POST + \epsilon_{it}$$

We found no treatment effect on *hproblem*, *hosp*, *gotcare*, but found a small treatment effect on *hpercepd*. People were 4% more likely to feel healthy.

```
# DID means hproblem
D <- pmb_data %>% group_by(post, pmb) %>% summarize(hproblem = mean(hproblem, na.rm = TRUE))
```

hproblem

```
## `summarise()` has grouped output by 'post'. You can override using the
## `.groups` argument.
```

```
# ATT hproblem
(D$hproblem[D$pmb & D$post] - D$hproblem[D$pmb & !D$post]) -
  (D$hproblem[!D$pmb & D$post] - D$hproblem[!D$pmb & !D$post])
```

```
## [1] 0.004097167
```

```r
# DID model hproblem
model1 <- lm(hproblem ~ pmb + pmb * post + post + cantcode2, data = pmb_data)
model1 <- coeftest(model1, vcov = vcovCL, cluster = ~cantcode2)
model1[which(!startsWith(row.names(model1), 'cantcode2')), ]
```

```
##                      Estimate Std. Error    t value     Pr(>|t|)
## (Intercept)       0.434208575 0.01020421 42.5519046 3.169053e-92
## pmbTRUE           0.036627930 0.01356970  2.6992443 7.655944e-03
## postTRUE         -0.202713585 0.02040842 -9.9328406 1.334343e-18
## pmbTRUE:postTRUE  0.004097167 0.02713940  0.1509675 8.801814e-01
```

```r
# DID model hosp
model2 <- lm(hosp ~ pmb + pmb * post + post + cantcode2, data = pmb_data)
model2 <- coeftest(model2, vcov = vcovCL, cluster = ~cantcode2)
model2[which(!startsWith(row.names(model2), 'cantcode2')), ]
```

hosp

```
##                      Estimate  Std. Error    t value     Pr(>|t|)
## (Intercept)       0.047261170 0.002142193 22.0620498 1.242019e-51
## pmbTRUE          -0.010893886 0.002554982 -4.2637823 3.335535e-05
## postTRUE          0.002665155 0.004284386  0.6220623 5.347392e-01
## pmbTRUE:postTRUE -0.002603302 0.005109963 -0.5094560 6.110972e-01
```

```r
# DID model gotcare
model3 <- lm(hosp ~ pmb + pmb * post + post + cantcode2, data = pmb_data)
model3 <- coeftest(model3, vcov = vcovCL, cluster = ~cantcode2)
model3[which(!startsWith(row.names(model3), 'cantcode2')), ]
```

gotcare

```
##                      Estimate  Std. Error    t value     Pr(>|t|)
## (Intercept)       0.047261170 0.002142193 22.0620498 1.242019e-51
## pmbTRUE          -0.010893886 0.002554982 -4.2637823 3.335535e-05
## postTRUE          0.002665155 0.004284386  0.6220623 5.347392e-01
## pmbTRUE:postTRUE -0.002603302 0.005109963 -0.5094560 6.110972e-01
```

```r
# DID model hpercepd
model4 <- lm(hpercepd ~ pmb + pmb * post + post + cantcode2, data = pmb_data)
model4 <- coeftest(model4, vcov = vcovCL, cluster = ~cantcode2)
model4[which(!startsWith(row.names(model4), 'cantcode2')), ]
```

hpercepd

```
##                     Estimate  Std. Error    t value     Pr(>|t|)
## (Intercept)       0.30261328 0.008570332  35.309398 6.495045e-80
## pmbTRUE          -0.19730180 0.011462962 -17.212113 5.251911e-39
## postTRUE          0.01872587 0.017140665   1.092482 2.761760e-01
## pmbTRUE:postTRUE  0.03917922 0.022925924   1.708948 8.929623e-02
```

Second DID specification model results

$$Y_{it} = \alpha_0 + \beta_1 pweekscov_{it} + \beta_2 POST + \beta_3 pweekscov_{it} \times POST + \epsilon_{it}$$

We found no treatment effect on any outcome variable *hproblem*, *hosp*, *gotcare*, and *hpercepd*.

```
# DID model hproblem vs pweekscov
model11 <- lm(hproblem ~ pweekscov + pweekscov * post + post + cantcode2, data = pmb_data)
model11 <- coeftest(model11, vcov = vcovCL, cluster = ~cantcode2)
model11[which(!startsWith(row.names(model11), 'cantcode2')), ]
```

hproblem

```
##                       Estimate  Std. Error     t value      Pr(>|t|)
## (Intercept)         0.43627443 0.008832029  49.3968534 2.319510e-102
## pweekscov           0.10300851 0.026328842   3.9123828  1.322262e-04
## postTRUE           -0.20684529 0.017664057 -11.7099535  1.435116e-23
## pweekscov:postTRUE  0.02706435 0.052657683   0.5139678  6.079459e-01
```

```
# DID model hosp vs pweekscov
model22 <- lm(hosp ~ pweekscov + pweekscov * post + post + cantcode2, data = pmb_data)
model22 <- coeftest(model22, vcov = vcovCL, cluster = ~cantcode2)
model22[which(!startsWith(row.names(model22), 'cantcode2')), ]
```

hosp

```
##                        Estimate  Std. Error    t value     Pr(>|t|)
## (Intercept)         0.047587270 0.001740282 27.3445686 6.001201e-64
## pweekscov          -0.034845522 0.004238438 -8.2213124 5.093436e-14
## postTRUE            0.002012955 0.003480565  0.5783415 5.638035e-01
## pweekscov:postTRUE -0.003804526 0.008476876 -0.4488122 6.541421e-01
```

```
model33 <- lm(gotcare ~ pweekscov + pweekscov * post + post + cantcode2, data = pmb_data)
model33 <- coeftest(model33, vcov = vcovCL, cluster = ~cantcode2)
model33[which(!startsWith(row.names(model33), 'cantcode2')), ]
```

gotcare

```
##                        Estimate Std. Error     t value      Pr(>|t|)
## (Intercept)          0.58253467 0.01115654  52.2146525 3.488912e-106
## pweekscov           -0.42007937 0.02948243 -14.2484629  9.130261e-31
## postTRUE             0.18819402 0.02231307   8.4342494  1.423634e-14
## pweekscov:postTRUE -0.02072332 0.05896487  -0.3514521  7.256870e-01
```

```
model44 <- lm(hpercepd ~ pweekscov + pweekscov * post + post + cantcode2, data = pmb_data)
model44 <- coeftest(model44, vcov = vcovCL, cluster = ~cantcode2)
model44[which(!startsWith(row.names(model44), 'cantcode2')), ]
```

hpercepd

```
##                        Estimate  Std. Error     t value     Pr(>|t|)
## (Intercept)          0.29427259 0.007561548  38.9169777 2.823460e-86
## pweekscov           -0.55008660 0.021611594 -25.4533105 1.038684e-59
## postTRUE             0.03540726 0.015123096   2.3412704 2.038485e-02
## pweekscov:postTRUE  0.02920286 0.043223187   0.6756294 5.001994e-01
```

Apendix: Data wrangling

```
if(!require(tidyverse)) install.packages("tidyverse", repos = "http://cran.us.r-project.org")

# -------------------------------------------------- #
# (1) canton/city data 2018
# -------------------------------------------------- #

ensanut_2018 <-
  mutate(ensanut_2018,

         # canton code
         cantcode2 = substr(upm,1,4),

         # survey date
         esdate = paste(fecha_mes, fecha_dia, fecha_anio, sep = "-"),
         esdate = as.Date(esdate, format = "%m-%d-%Y"),

         # indigenous
         indrace_2018 = etnia == 1)

# -------------------------------------------------- #
# (2) canton/city data 2012
# -------------------------------------------------- #

ensanut_2012 <-
  mutate(ensanut_2012,

         # canton code
         cantcode2 = case_when(prov >= 10 ~ substr(ciudad,1,4),
                               prov < 10 ~ paste(0, substr(ciudad,1,3), sep = "")),
```

```r
         # indigenous
         indrace_2012 = gr_etn == 1)


# ------------------------------------------------- #
# (3) outcome variables 2018
# ------------------------------------------------- #

ensanut_2018 <-
  mutate(ensanut_2018,

         # had health problem in past 30 days
         hproblem_2018 = f1_s4_2 == 1,

         # hospitalization in past 12 months
         hosp_2018 = f1_s4_54 == 1,

         # received care to address health problems in past 30 days
         gotcare_2018 = case_when(is.na(f1_s4_6) ~ NA,
                                 f1_s4_6 %in% c(1,2,4) ~ TRUE,
                                 TRUE ~ FALSE),

         # received preventive care services in past 30 days
         prev_2018 = f1_s4_41 == 1,

         # health status perception
         hpercep_2018 = case_when(f1_s4_58 == 1 ~ 5, f1_s4_58 == 2 ~ 4,
                                 f1_s4_58 == 4 ~ 2, f1_s4_58 == 5 ~ 1, TRUE ~ 3),

         # health status perception dummy
         hpercepd_2018 = hpercep_2018 %in% c(4,5))


pmb_data_2018 <- ensanut_2018 %>% group_by(cantcode2) %>%
  summarize(esdate = median(as.numeric(esdate), na.rm = TRUE),
            n_2018 = n(),
            hproblem_2018 = weighted.mean(hproblem_2018, w = fexp, na.rm = TRUE),
            hosp_2018 = weighted.mean(hosp_2018, w = fexp, na.rm = TRUE),
            gotcare_2018 = weighted.mean(gotcare_2018, w = fexp, na.rm = TRUE),
            prev_2018 = weighted.mean(prev_2018, w = fexp, na.rm = TRUE),
            hpercepd_2018 = weighted.mean(hpercepd_2018, w = fexp, na.rm = TRUE),
            indrace_2018 = weighted.mean(indrace_2018, w = fexp, na.rm = TRUE))

# cantons with high indigenous concentration
pmb_data_2018$ind_2018 <- pmb_data_2018$indrace_2018 > mean(pmb_data_2018$indrace_2018, na.rm = TRUE)

# ------------------------------------------------- #
# (1) outcome variables 2012
# ------------------------------------------------- #

ensanut_2012 <-
  mutate(ensanut_2012,
```

```
         # had health problem in past 30 days
         hproblem_2012 = ps02 == 1,

         # hospitalization in past 12 months
         hosp_2012 = ps55 == 1,

         # received care to address health problems in past 30 days
         gotcare_2012 = case_when(is.na(ps06) ~ NA,
                                  ps06 %in% c(1,2,4) ~ TRUE,
                                  TRUE ~ FALSE),

         # received preventive care services in past 30 days
         prev_2012 = ps40 == 1,

         # health status perception
         hpercep_2012 = case_when(ps71 == 1 ~ 5, ps71 == 2 ~ 4,
                                  ps71 == 4 ~ 2, ps71 == 5 ~ 1, TRUE ~ 3),

         # health status perception binary
         hpercepd_2012 = hpercep_2012 %in% c(4,5))


pmb_data_2012 <- ensanut_2012 %>% group_by(cantcode2) %>%
  summarize(n_2012 = n(),
            hproblem_2012 = weighted.mean(hproblem_2012, w = pw, na.rm = TRUE),
            hosp_2012 = weighted.mean(hosp_2012, w = pw, na.rm = TRUE),
            gotcare_2012 = weighted.mean(gotcare_2012, w = pw, na.rm = TRUE),
            prev_2012 = weighted.mean(prev_2012, w = pw, na.rm = TRUE),
            hpercepd_2012 = weighted.mean(hpercepd_2012, w = pw, na.rm = TRUE),
            indrace_2012 = weighted.mean(indrace_2012, w = pw, na.rm = TRUE))

# cantons with high indigenous concentration
pmb_data_2012$ind_2012 <- pmb_data_2012$indrace_2012 > mean(pmb_data_2012$indrace_2012, na.rm = TRUE)

# -------------------------------------------------- #
# (1) PMB dates
# -------------------------------------------------- #

pmb_cantons_date <-
  mutate(pmb_cantons_date,

         # canton code
         cantcode2 = case_when(province_code < 10 & canton_code < 10 ~
                                 paste(0, province_code, 0, canton_code, sep = ""),
                               province_code < 10 & canton_code >= 10 ~
                                 paste(0, province_code, canton_code, sep = ""),
                               province_code >= 10 & canton_code < 10 ~
                                 paste(province_code, 0, canton_code, sep = ""),
                               province_code >= 10 & canton_code >= 10 ~
                                 paste(province_code, canton_code, sep = "")),

         # PMB start date
         pmb_start_date = as.Date(pmb_start_date, format = "%m/%d/%Y"),
```

```
        pmb_start_date = as.integer(pmb_start_date)) %>%
  select(cantcode2, pmb_start_date)

# removes duplicates
pmb_cantons_date <- pmb_cantons_date[!duplicated(pmb_cantons_date),]

# -------------------------------------------------- #
# (4) treatment variables
# -------------------------------------------------- #

pmb_data <- left_join(pmb_data_2018, pmb_data_2012, by = "cantcode2") %>%
  left_join(pmb_cantons_date, by = "cantcode2")

# number of weeks a given canton was receiving the MBP program before the survey
pmb_data$weeks_ensanut18 <- (pmb_data$esdate - pmb_data$pmb_start_date) / (365/52)

# PMB dummy
pmb_data$pmb <- pmb_data$weeks_ensanut18 > 0
mean(pmb_data$pmb, na.rm = TRUE) # percentage of treatment cantons

# maximum number of covered weeks by canton
# PMB program officially started on August 30 2017
weeks_possible_ensanut18 <- (max(pmb_data$esdate) - as.integer(as.Date("2017-08-30"))) / (365/52)

pmb_data$pweekscov <- pmb_data$weeks_ensanut18 / weeks_possible_ensanut18
pmb_data$pweekscov <- ifelse(pmb_data$pweekscov < 0, 0, pmb_data$pweekscov)

pmb_data <- pmb_data[complete.cases(pmb_data), ] # remove incomplete cases

# final dataset
pmb_data <- pmb_data %>% gather("var", "value", n_2018: ind_2012, convert = TRUE) %>%
  separate(var, c("var", "year")) %>% spread(var, value)

# post variable
pmb_data$post <- pmb_data$year == 2018

saveRDS(pmb_data, "data/pmb_data.rds"
```

Referenes

Angrist, J. D., & Pischke, J.-S. (2009). Mostly harmless econometrics: An empiricist's companion. Princeton, NJ: Princeton University Press.

Angrist, J. D., & Pischke, J.-S. (2015). Mastering 'Metrics: The Path from Cause to Effect. Princeton, NJ: Princeton University Press.

Athey, S., & Imbens, G. W. (2022). Design-based analysis in Difference-In-Differences settings with staggered adoption. Journal of Econometrics, 226(1), 62–79. https://doi.org/10.1016/j.jeconom.2020.10.012