11-23-2019

Concept Note: Mexico

Danae Hernandez
University of California, Santa Barbara

Alejandro López-Feldman
Tecnológico de Monterrey

Fernanda Márquez-Padilla
Centro de Investigación y Docencia Económicas

Concept note: Mexico

Danae Hernandez (UCSB)
Alejandro López-Feldman (Tec)
Fernanda Márquez-Padilla (CIDE)

Forest Collaborative
EfD

Health data

- Individual medical records and emergency room visits from the Mexican National Social Security Institute (IMSS).
- IMSS is the largest health care provider in Mexico (approximately, 58% of the Mexican population).
- The data comes from individual electronic records generated by clinics and hospitals.
- We have access to 1,120/1,506 family units, 246/248 second level clinics, and 37/37 speciality care units.

More detail on the data

- We have data on routine doctor visits and emergency visits for children and young adults in ages 0 to 25.
- Weight, height, clinical diagnostic (International Statistical Classification of Diseases/ ICD-10)

More detail on the data

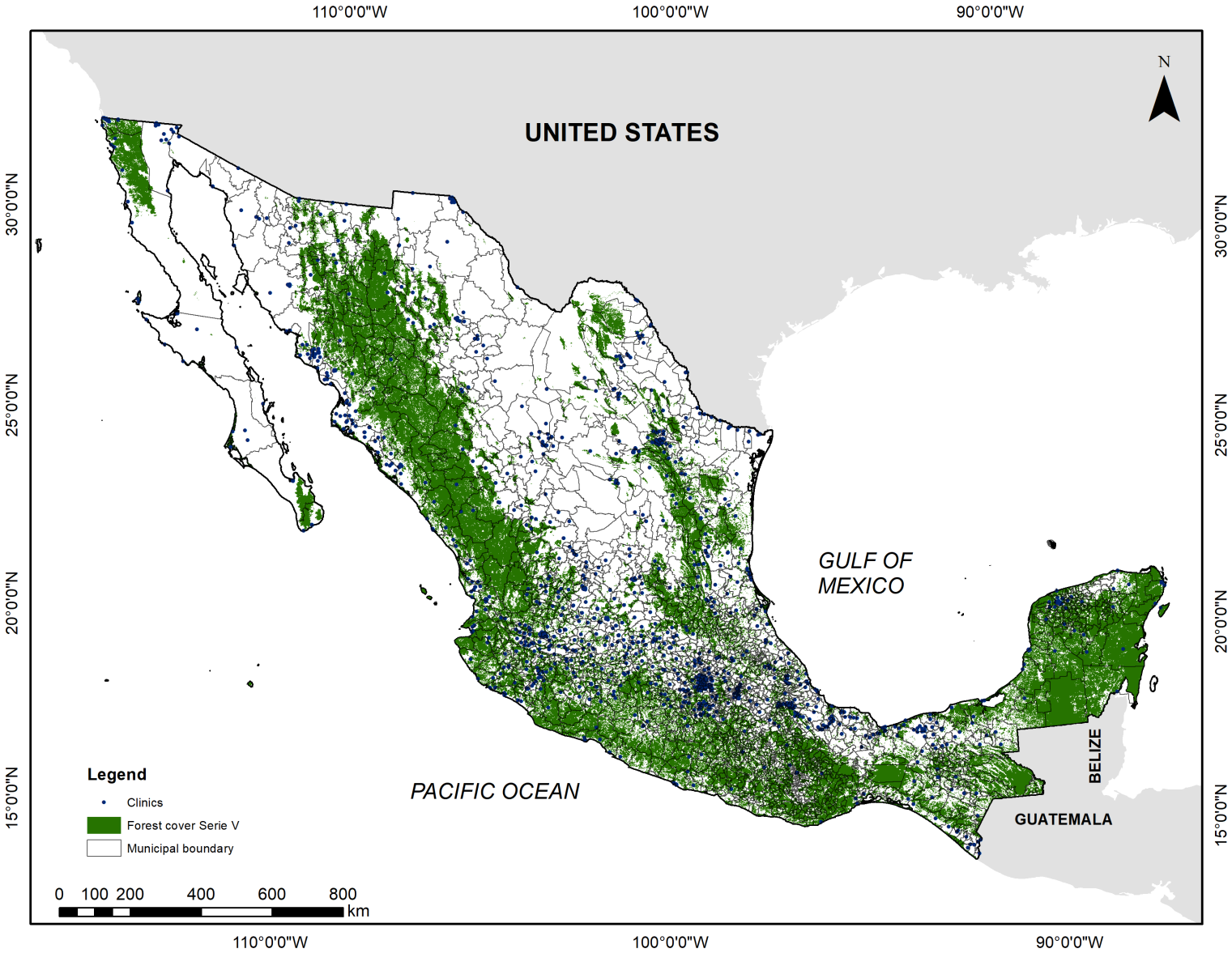
- We have data on routine doctor visits and emergency visits for children and young adults in ages 0 to 25.
- Weight, height, clinical diagnostic (International Statistical Classification of Diseases/ ICD-10)

Table 1: Descriptive statistics
Full sample

	(1)		
	mean	sd	count
Age	12.301	(8.774)	130,453,619
Male	0.442	(0.497)	130,453,582
Weight (kg)	42.370	(27.067)	108,430,906
Height (m)	1.314	(0.364)	106,956,650
Body Mass Index	26.062	(5.148)	40,494,074
BMI-for-age (z-scores)	0.492	(1.473)	64,593,105
ICD A, dengue	0.002	(0.046)	130,453,619
ICD A, chikungunya	0.001	(0.023)	130,453,619
ICD A, diarrhea	0.034	(0.180)	130,453,619
ICD A: infectious	0.051	(0.219)	130,453,619
ICD J: respiratory	0.279	(0.448)	130,453,619
ICD K: digestive	0.100	(0.300)	130,453,619
Visit type: emergency	0.194	(0.396)	130,453,619
Visit type: family doctor	0.715	(0.451)	130,453,619

Location of IMSS clinics in the sample

Location of IMSS clinics in the sample



Pros and cons of the data

Pros and cons of the data

- Pros:
  - High quality administrative data.
  - Large sample for the period 2011-2017.
  - The panel structure of the data allows us to follow patients in time.
  - More than 130 million observations (clinic visits) from approximately 18 million patients.

Pros and cons of the data

- Pros:
  - High quality administrative data.
  - Large sample for the period 2011-2017.
  - The panel structure of the data allows us to follow patients in time.
  - More than 130 million observations (clinic visits) from approximately 18 million patients.
- Cons:
  - Data comes from electronic records and most rural clinics do not use them.
  - Therefore, our sample does not cover rural remote areas.
  - 25% of the family clinics in Mexico are not in the sample.
  - Haven't checked for differences in characteristics of missing clinics.

Some ideas to look at forests and health

Some ideas to look at forests and health

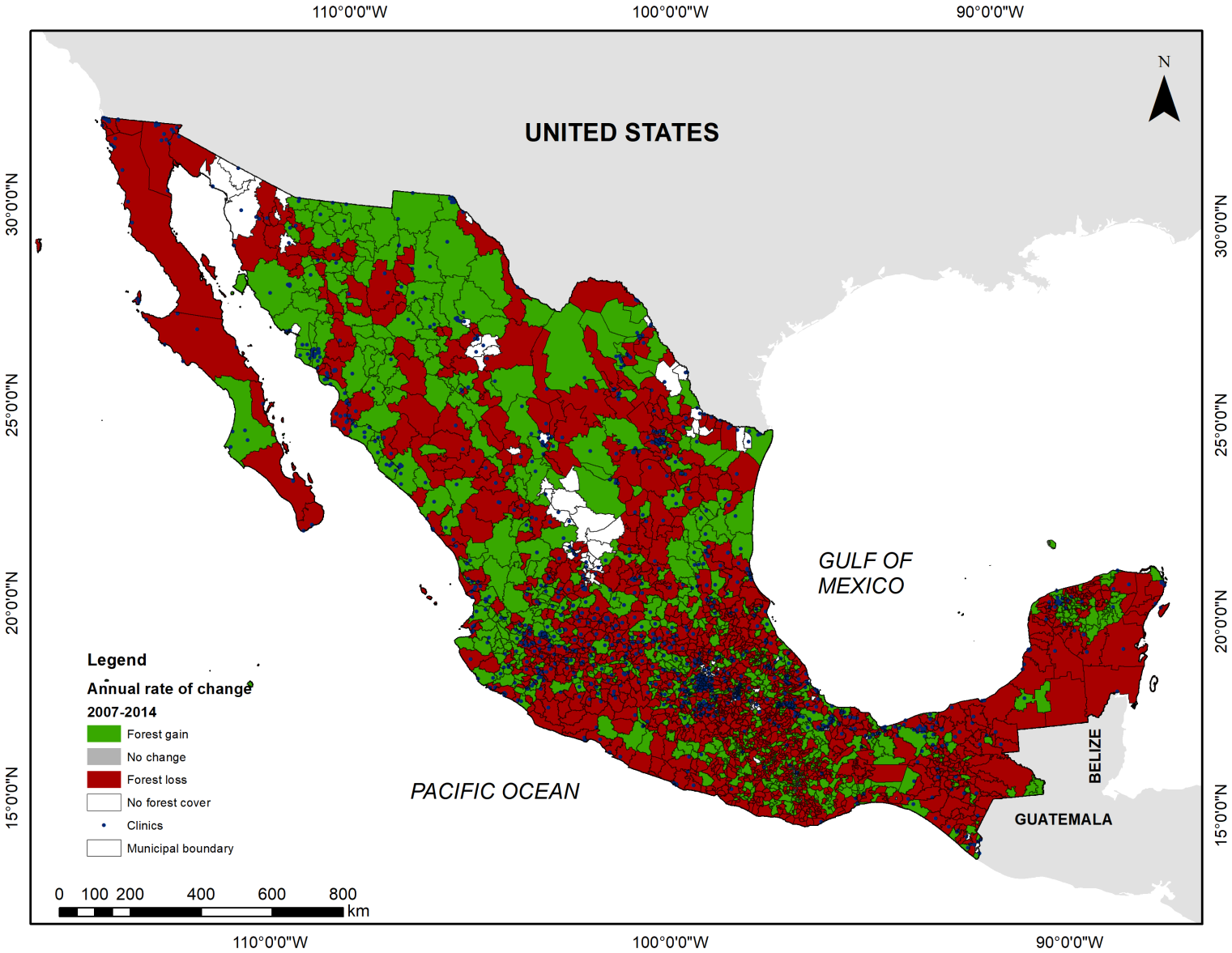
- Look at effects of forest loss (forest cover) on multiple outcomes (diarrhea, dengue, etc.)
- Create 5-km (10-km) buffer around clinic location to allocate forest loss (forest cover) to the individuals using that clinic.

Some ideas to look at forests and health

- Look at effects of forest loss (forest cover) on multiple outcomes (diarrhea, dengue, etc.)
- Create 5-km (10-km) buffer around clinic location to allocate forest loss (forest cover) to the individuals using that clinic.
- Use Hansen data (annual variation) and INEGI data (more detail on forest type).

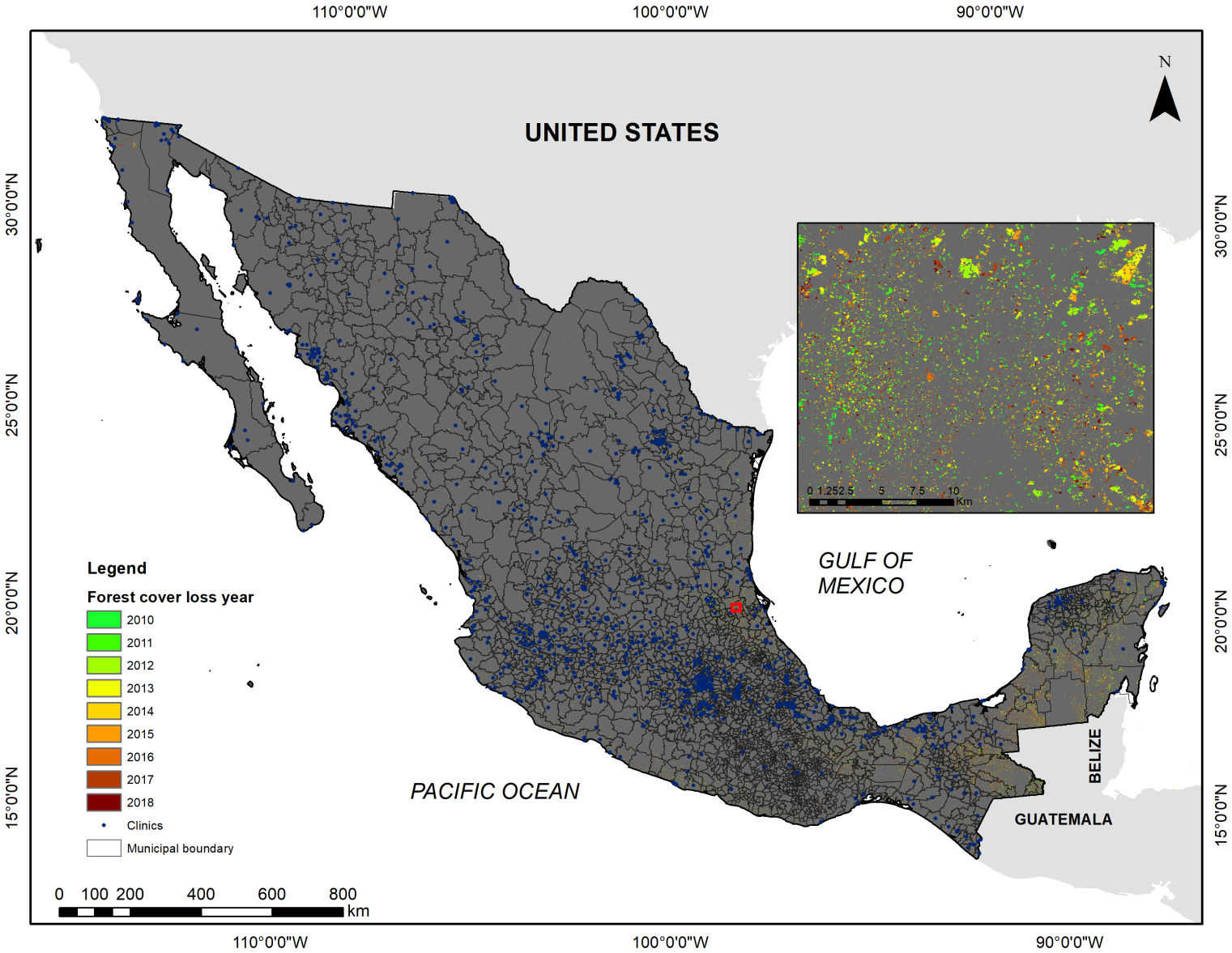
Forest change and location of clinics

Forest change and location of clinics



Forest loss by year

Forest loss by year



Deforestation and health (I)

Deforestation and health (I)

Table 2: Descriptive statistics by deforestation, 2001-2017
Random sample (5km buffer)

	(1) Above median	(2) Below median	(3) Difference
Age	9.300	9.247	***
	(6.033)	(5.987)	
Male	0.491	0.494	**
	(0.500)	(0.500)	
Weight (kg)	35.83	35.83	
	(22.29)	(22.05)	
Height (m)	1.288	1.293	***
	(0.294)	(0.297)	
ICD A, dengue	0.00287	0.000992	***
	(0.0535)	(0.0315)	
ICD A, chikungunya	0.000654	0.0000851	***
	(0.0256)	(0.00922)	
ICD A, diarrhea	0.0429	0.0403	***
	(0.203)	(0.197)	
ICD A: infectious	0.0580	0.0511	***
	(0.234)	(0.220)	
ICD J: respiratory	0.360	0.376	***
	(0.480)	(0.484)	
ICD K: digestive	0.109	0.110	*
	(0.311)	(0.313)	
Observations	781,181	599,539	1,380,720

mean coefficients; sd in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Deforestation and health (II)

Deforestation and health (II)

Table 3: Descriptive statistics by deforestation, 2010-2017
Random sample (5km buffer)

	(1) Above median	(2) Below median	(3) Difference
Age	9.275	9.279	
	(6.027)	(5.995)	
Male	0.491	0.495	***
	(0.500)	(0.500)	
Weight (kg)	35.66	36.05	***
	(22.22)	(22.14)	
Height (m)	1.286	1.296	***
	(0.295)	(0.296)	
ICD A, dengue	0.00305	0.000789	***
	(0.0551)	(0.0281)	
ICD A, chikungunya	0.000693	0.0000428	***
	(0.0263)	(0.00654)	
ICD A, diarrhea	0.0430	0.0402	***
	(0.203)	(0.196)	
ICD A: infectious	0.0584	0.0507	***
	(0.234)	(0.219)	
ICD J: respiratory	0.362	0.374	***
	(0.481)	(0.484)	
ICD K: digestive	0.108	0.110	**
	(0.311)	(0.313)	
Observations	773,691	607,029	1,380,720

mean coefficients; sd in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Deforestation and health (III)

Deforestation and health (III)

Table 6: Descriptive statistics by deforestation, age 0-5, 2001-2017
Random sample (5km buffer)

	(1) Above median	(2) Below median	(3) Difference
Age	2.820	2.783	***
	(1.609)	(1.617)	
Male	0.539	0.538	
	(0.498)	(0.499)	
Weight (kg)	15.14	15.08	**
	(6.804)	(6.645)	
Height (m)	0.954	0.953	*
	(0.170)	(0.172)	
ICD A, dengue	0.00105	0.000240	***
	(0.0324)	(0.0155)	
ICD A, chikungunya	0.000243	0.0000204	***
	(0.0156)	(0.00452)	
ICD A, diarrhea	0.0596	0.0561	***
	(0.237)	(0.230)	
ICD A: infectious	0.0689	0.0632	***
	(0.253)	(0.243)	
ICD J: respiratory	0.496	0.518	***
	(0.500)	(0.500)	
ICD K: digestive	0.0592	0.0577	*
	(0.236)	(0.233)	
Observations	255,397	196,135	451,532

mean coefficients; sd in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Forest cover and health (I)

Forest cover and health (I)

Table 4: Descriptive statistics by forest cover, 2011
Random sample (5km buffer)

	(1) Above median	(2) Below median	(3) Difference
Age	9.171	9.319	***
	(5.907)	(6.036)	
Male	0.488	0.495	***
	(0.500)	(0.500)	
Weight (kg)	35.53	35.98	***
	(21.93)	(22.21)	
Height (m)	1.288	1.292	***
	(0.293)	(0.296)	
ICD A, dengue	0.00167	0.00201	***
	(0.0408)	(0.0448)	
ICD A, chikungunya	0.000305	0.000477	***
	(0.0175)	(0.0218)	
ICD A, diarrhea	0.0447	0.0406	***
	(0.207)	(0.197)	
ICD A: infectious	0.0579	0.0537	***
	(0.234)	(0.225)	
ICD J: respiratory	0.382	0.361	***
	(0.486)	(0.480)	
ICD K: digestive	0.112	0.109	***
	(0.315)	(0.312)	
Observations	384,103	1,183,636	1,567,739

mean coefficients; sd in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Forest cover and health (II)

Forest cover and health (II)

Table 5: Descriptive statistics by forest cover, 2011, age 0-5
Random sample (5km buffer)

	(1) Above median	(2) Below median	(3) Difference
Age	2.826	2.804	***
	(1.620)	(1.609)	
Male	0.532	0.543	***
	(0.499)	(0.498)	
Weight (kg)	15.17	15.10	**
	(6.688)	(6.744)	
Height (m)	0.954	0.954	
	(0.172)	(0.170)	
ICD A, dengue	0.000524	0.000683	
	(0.0229)	(0.0261)	
ICD A, chikungunya	0.000111	0.000174	
	(0.0105)	(0.0132)	
ICD A, diarrhea	0.0614	0.0570	***
	(0.240)	(0.232)	
ICD A: infectious	0.0701	0.0652	***
	(0.255)	(0.247)	
ICD J: respiratory	0.517	0.500	***
	(0.500)	(0.500)	
ICD K: digestive	0.0613	0.0585	***
	(0.240)	(0.235)	
Observations	125,978	385,211	511,189

mean coefficients; sd in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A simple empirical strategy to start

- Estimate a model like the following:

$$Y_{it} = \beta_0 + \beta_1 F_{it} + \beta_2 T_{it} + \beta_3 P_{it} + \mu_i + u_{it}$$

where Y_{it} is a health outcome for individual i in period t, F_{it} is the forest related variable (e.g., forest loss, forest cover), T_{it} and P_{it} are temperature and precipitation, respectively, and μ_i are individual level fixed effects.

A simple empirical strategy to start

- Estimate a model like the following:

$$Y_{it} = \beta_0 + \beta_1 F_{it} + \beta_2 T_{it} + \beta_3 P_{it} + \mu_i + u_{it}$$

where Y_{it} is a health outcome for individual i in period t, F_{it} is the forest related variable (e.g., forest loss, forest cover), T_{it} and P_{it} are temperature and precipitation, respectively, and μ_i are individual level fixed effects.

Things to consider

- ... hard to look at behavioral responses with this data.
- Maybe use municipal data to look at distributional effects.
- No individual level socioeconomic arguably is included in the fixed effect...

Things to consider

- ... hard to look at behavioral responses with this data.
- Maybe use municipal data to look at distributional effects.
- No individual level socioeconomic arguably is included in the fixed effect...
- ¿Some specific public policy to look at? Maybe at a regional level?

Things to consider

- ... hard to look at behavioral responses with this data.
- Maybe use municipal data to look at distributional effects.
- No individual level socioeconomic arguably is included in the fixed effect...
- ¿Some specific public policy to look at? Maybe at a regional level?
- Maybe use only data for family units.
- Estimate for full sample and by age groups.
- Can use deforestation lagged.
- Can separate by type of forest.

Thanks!