IDB WORKING PAPER SERIES No. IDB-WP-497

Assessing the Effects of Climate and Socioeconomic Factors on Vulnerability to Vector-Borne Diseases in Latin America

Rodrigo García Ayala
Andrés Estrugo

July 2014

Inter-American Development Bank
Department of Research and Chief Economist

Assessing the Effects of Climate and Socioeconomic Factors on Vulnerability to Vector-Borne Diseases in Latin America

Rodrigo García Ayala
Andrés Estrugo

Soluciones en Economía Aplicada (SEA)

Inter-American Development Bank
2014

Cataloging-in-Publication data provided by the
Inter-American Development Bank
Felipe Herrera Library

García Ayala, Rodrigo.
Assessing the effects of climate and socioeconomic factors on vulnerability to vector-borne diseases in Latin America / Rodrigo García Ayala, Andrés Estrugo.
p. cm. — (IDB Working Paper Series ; 497)
Includes bibliographic references.
1. Climatic changes—Health aspects—Latin America. 2. Dengue—Social aspects —Latin America. 3. Dengue—Economic aspects—Latin America. 4. Dengue—Climatic factors—Latin America. I. Estrugo, Andrés. II. Inter-American Development Bank. Department of Research and Chief Economist. III. Title. IV. Series.
IDB-WP-497

http://www.iadb.org

Abstract[1]

Climate change is imposing a large burden on the most vulnerable populations, particularly in the developing world. Establishing consistent causal relationships, however, is difficult because a multiplicity of climatic, economic and socio-demographic elements are combined to create the conditions for an outbreak of vector-borne disease. Based on a two-step procedure, this paper presents and tests an approach to estimating the effects of epidemic outbreaks on health vulnerability. The model proposed is empirically tested for five countries in Latin America where dengue is a national health priority. Using data from national censuses, satellite climate information and data from a newly developed disease outbreak surveillance online platform, the paper finds that climate has non-negligible effects on health vulnerability. The evidence found and the vulnerability index constructed can be used to analyze the main determinants of vulnerability in order to address policy concerns.

JEL classifications: D04, Q51, Q54, R58
Keywords: Vulnerability, Climate change, Socioeconomic conditions, Vector-borne diseases, Factor analysis

[1] The corresponding author is Rodrigo García Ayala (email: rodrigo.garcia@sea-cg.com). The authors are affiliated with Soluciones en Economía Aplicada (SEA), Cochabamba, Bolivia. The authors thank the Inter-American Development Bank for the grant offered to carry out this study. Inés Lucía Tejada M. provided valuable research assistance, and Henry Durán and Marcos Andrade supported the retrieval and processing of climatic and geo-referenced information. García also acknowledges Health Maps and IPUMS for granting access to their databases. The usual disclaimers regarding responsibility of mistakes and errors apply. The authors additionally acknowledge valuable comments by Sebastián Miller, Nora Libertun, Jorge Agüero and the participants in the Health Impacts of Climate Change in Latin America and the Caribbean Seminar held at the Inter-American Development Bank in October, 2013.

1. Introduction

Climate change is imposing a burden that is unevenly distributed across populations, and the most vulnerable tend to bear a relative larger share of this burden. In recent years, the frequency and intensity of extreme climatic episodes has increased, challenging communities' ability to cope with the adverse effects of those events. Naturally, the most vulnerable populations are in worse positions to effectively face these challenges to the extent that they have a limited ability to withstand the impact of extreme weather events and adapt to changing conditions. Therefore, regardless of how the causes of climate change are assessed, this new scenario calls for policy strategies aimed at increasing the resilience of local communities and thereby reducing the burden of the most vulnerable populations.

Latin America's adaptation is complicated by an unprecedented urbanization process that has made it the second most urbanized region in the world, only behind North America (IDB, 2011). Most intermediate Latin American and Caribbean cities have experienced impressive growth rates in their urban populations in recent years, a trend that is very likely continue in the years ahead. This has placed additional pressure on already-deficient provision of basic services and increased the vulnerability of newcomers.

A particular case of vulnerability in this context is the emergence of vector-borne diseases. Dengue, malaria, leishmaniasis, Chagas disease and tick-borne diseases are endemic in many countries in Latin America and the Caribbean, and episodic outbreaks are thus common in the region. Figures from Health Map,[2] a web-based platform for epidemic surveillance, indicate that in 2013 more than 5,236 vector-borne outbreak alerts (5,207 for dengue) took place in 17 countries where these diseases are endemic. Vulnerability to vector-borne diseases can clearly have a significant impact on Latin American citizens' quality of life and their productivity. Given the large share of dengue outbreaks, this paper's analysis is focused on that disease.

There is an increasing interest in both public policy and academic audiences in estimating the impacts of extreme events on health outcomes. This being a relatively new area, much of the research has concentrated on the impacts of extreme temperatures on mortality, agriculture and industrial outcomes, and even social conflict (Deschênes, 2012, and Dell, Jones and Olken, 2013, offer a comprehensive literature review of the effects of extreme events on health outcomes).

[2] http://healthmap.org

Less research has been conducted on the effects of vector-borne diseases, given climate and social conditions, on health vulnerability. This paper seeks to bridge the gap in this regard.

Vulnerability has received significant attention in the literature. The epidemiological research has focused on the impacts of global climate change and population characteristics on both endemic and epidemic vector-borne diseases. Sutherst (2004) offers a summary of the complex interrelationship between global climate change, localized weather variability, biological adaptability, population socioeconomic characteristics, and the prevention and management of vector-borne diseases. Sutherst emphasizes the role of social and economic characteristics in the prevention and management of vector-borne diseases. Increasing standards of living in developing countries and proper allocation of resources are seen as vital tools in the prevention of vector-borne diseases. Haines and Patz (2004) also note the direct relationship between extreme weather patterns and the increased incidence of vector-borne diseases, while Haines et al. (2006) provide evidence that negative health effects of global climate change are more prevalent in low-income countries.

Furthermore, policy-driven studies such as those of Ebi, Kovats and Menne (2006) provide a framework for analyzing vulnerability to vector-borne diseases. An essential component of the policy framework is the proper assessment of the potential impacts of climate change and variability on different socioeconomic sectors within a population. Policy to promote prevention or management of vector-borne diseases could then be tailored based on this assessment. The statistical modeling approach provided in this paper therefore combines climate, social, and health data to provide a quantitative measurement of vulnerability to vector-borne diseases and extreme climate conditions.

A widely accepted notion of vulnerability defines it as the product of exposure, sensitivity and capacity to adaptation once the extreme event takes place (Few, 2007; Miller, Yoon and Yu, 2013; WHO, 2003). Exposure usually refers to the probability of an adverse event taking place, as well as the intensity of such an event. In the particular case of vector-borne diseases, exposure will be the result of both climatic and socioeconomic conditions in a given region. Sensitivity is defined as the impact the adverse event can have on the unit of analysis (a city from here on). Adaptation refers to the measures the city can take to neutralize the negative impacts. To a large extent, sensitivity and adaptation are two sides of the same coin. Both will be the result of a multitude of factors, including publicly provided facilities, as well as assets and strategies individual households and communities within a city have on hand to cope with the adverse effects. In this paper we propose to use this notion, along with quasi-experimental impact evaluation techniques, to estimate the impacts of climatic extreme events and socioeconomic conditions on health outcomes. Throughout the paper, it is assumed that vulnerability has a first order negative impact on health outcomes (i.e., higher levels of vulnerability will translate into poorer health outcomes), an assumption empirically tested in several papers (Molina, 2009; Tarazona and Gallegos, 2011).

Establishing a consistent causal relationship from these variables to health outcomes requires taking into account that a multiplicity of climatic, economic and socio-demographic elements are combined to create the conditions for the outbreak of diseases, and the subsequent effect on population wellbeing. In this paper two direct effects through which climatic and socioeconomic conditions can have an impact on vulnerability are estimated.

Vector-borne diseases are a particularly good example of these complex relationships. Along with a city's economic development and basic service coverage, social institutions and households' capabilities, climatic and geographic conditions are to large extent determinants of vector populations and thus influence the likelihood of outbreaks. Temperature, precipitation, and humidity are critical to mosquito survival, reproduction, and development and thus influence mosquito presence and abundance (Githeko et al., 2000; Hii, 2013). Extreme weather events can directly affect the likelihood of epidemic outbreaks by enhancing the environmental conditions that allow mosquito populations to thrive. Thus, ceteris paribus, increases (within a range) in either (or both) precipitation and temperature are expected to have a positive impact on the expected likelihood of an outbreak.

Climate change may shift the pattern distribution of vulnerability across regions in Latin America. Dengue and malaria require specific climatic conditions for growth (Githeko et al., 2000). Increased levels of rainfalls, and variations in the temperature regimes across regions derived from climate change may condition the appearance of susceptible mosquito populations.

For instance, in countries where transmission does routinely occur, short-term changes in weather (temperature, precipitation, and humidity) are often correlated with higher levels of incidence of dengue,[3] the most prevalent vector-borne disease in the region. Climate change is likely to increase the frequency of such extreme short-term climatic events. Furthermore, climate

[3] http://www.cdc.gov/dengue/entomologyEcology/climate.html#climate

change may change the average levels of precipitation and temperature of certain regions, enabling the emergence of vector populations where few or no historical records of such diseases exist.

A growing literature has robustly demonstrated the direct effects of extreme climate conditions on health outcomes (see Deschênes, 2012, and Dell, Jones and Olken, 2013, for a review). Direct effects consist mainly of physical discomfort but can encompass serious illness and even death. These effects are generally exacerbated in children and the elderly, the populations most vulnerable to these diseases. Several papers have aimed to quantify these effects (Guerrero, 2013; Burgess et al., 2011; WHO, 2003; Githeko et al., 2000; Molina, 2009; Patz et al., 2005; Deschênes and Moretti, 2007).

A second set of indirect effects also takes place. Indirectly, being infected with malaria or dengue will entail loss of working days, with subsequent impact on household income-generating capabilities. This is the *income mechanism* explored in Burgess et al. (2011) and Guerrero (2013) for India and Mexico, respectively. In this paper we will not address effects of that type.

However, no epidemic outbreak takes place in a social vacuum. Urbanization can either create the conditions for epidemic outbreaks to thrive or halt their appearance, conditional on the quality of basic services and other types of basic and public health infrastructure (Friel et al., 2011; Kjellstrom et al., 2007; Vlahov et al., 2007). The availability of some public services (mainly sewage and water supply) varies considerably across cities, particularly in large and intermediate metropolitan areas, which already constitute a vast majority of urban settlements in Latin America (IDB, 2011; Carrera, 2013). Local imbalances in access to publicly provided services and in possession of assets increase the pressure on vulnerable populations, who are less capable of adapting to environmental threats stemming from climate change. This point can be particularly relevant in highly dense locations, where a relatively large portion of the local population is vulnerable.

In this regard, socioeconomic conditions have two direct effects on the adopted measure of health vulnerability. The first, as in the case of extreme climate, is directly linked to the likelihood of epidemic outbreaks in a given region. The degree and quality of health and infrastructure services are expected to have a negative first order effect on the likelihood of an outbreak.

The evidence found in the literature is robust with regard to this causal link. Nonetheless, some services will tend to have a larger impact on the likelihood of an outbreak than others; sewage, access to piped water, and provision of health services, among others, are expected to be better predictors of outbreaks than other services (Kjellstrom et al., 2007; Vlahov et al., 2007).

A second set of direct effects comes as a result of the definition of sensitivity used. As mentioned, sensitivity is the result of city-wide, community and household-level capabilities that enable individuals and households to cope with adverse events. Following the ideas presented in Cutter, Boruff and Shirley (2003) and IPCC (2007), this framework allows us to estimate an index composed of two different dimensions (a household-level factor and a community/city factor) that add up to represent the abstract concept of sensitivity. The index incorporates only socioeconomic information extracted from national census sources.

Bringing all these elements together, a means of accounting for both set of impacts—i.e., the effects of climatic extremes and socio-economic conditions on vector-borne disease vulnerability—is presented in this paper. Central to our estimation strategy is the notion that climatic covariates are strictly exogenous. We estimate the model presented in Section 3 for five Latin American countries where dengue is prevalent and represents a national health priority. Our analysis can be regarded as exploratory, as we aim to estimate the effects of structural factors on health vulnerability.

The research contributes to the existing literature in several regards. First, the scope of our approach allows us to obtain implications that are to a large extent externally valid, at least for Latin America, as more than 1,700 municipalities in five countries are included in the sample. As a result, by differentiating climatic from socioeconomic effects on vulnerability in a regional-scale study, it is possible to contribute to the discussion of appropriate policy measures to reduce vulnerability and increase community resilience to the effects of climate change.

Second, the empirical strategy adopted addresses causality issues. The internal validity of the study is guaranteed by two considerations. First, climate is assumed to cause (or at least contribute) to epidemic outbreaks. Where internal validity cannot be guaranteed, as is the case for socioeconomic conditions, differences in differences estimates are used to eliminate unobserved and assumed time-invariant sources of endogeneity. However, constructing a new sensitivity index that appropriately incorporates the effects of suffering a vector borne disease becomes mandatory in order to improve the conclusions from the analysis here presented.

Finally, to our knowledge this is one of the handful efforts to assess the impacts of climate extremes on health vulnerability to vector-borne diseases in Latin America (Pereda, 2014; Valencia, 2014).Very few studies are in a position to address the role of adaptation explicitly (Deschênes, 2012); here, adaptation is not explicitly considered, but rather an overall measure of sensitivity, which includes elements that can be regarded as adaptation enhancers.

The paper is divided in six parts. Section 2 describes the literature linking climatic and socioeconomic conditions to vector-borne diseases prevalence in Latin America, paying particular attention to dengue fever and malaria. In the third section the conceptual framework utilized to estimate the effects of climate change on vulnerability and health is introduced. Section 4 describes the data used. Section 5 presents some preliminary results. Finally, Section 6 concludes.

2. Vector-Borne Diseases and Extreme Climate

2.1. Situation in Latin America and the Caribbean

Dengue, malaria, leishmaniasis, Chagas disease and tick-borne diseases are endemic in several countries in Latin America and the Caribbean, with public health efforts most concentrated on dengue fever and malaria. Dengue fever is by far the most prevalent vector-borne disease in the region, and malaria the most menacing because of the large burden imposed on the infected population. The subsequent analysis and description will be entirely focused on dengue, however, because malaria in Latin America is less endemic than in other regions and manifests mainly in rural areas.

Table 1 presents a summary of data collected for fourteen countries in the region. The first striking fact is the large number of average cases of dengue in the region and the considerably higher incidence relative to malaria, reaching ratios larger than thousands in some countries. These large differences may be explained by malaria eradication efforts across the region and by dengue-transmitting mosquitoes' behavior and favorable habitat in growing urban settings.

Based on Table 1, it is possible to classify countries in the region into three broad categories based on the incidence and number of reported cases. The subsequent analysis, which attempts to obtain evidence on the impact climatic and socioeconomic conditions can have on the emergence of vector-borne diseases, will consider this basic classification along with other structural and institutional factors (such as GDP, quality of institutions, and geographic conditions. The three groups are the following:

1. High number of cases and relative low incidence: Brazil, Mexico.
2. Relative low number of cases but relative high incidence: Belize, Bolivia, Costa Rica, El Salvador, Nicaragua and Paraguay.
3. None of the above: Colombia, Honduras, Peru, Dominican Republic and Venezuela.

This rough classification provided the basis for selecting the five countries included in the empirical analysis. Brazil and Mexico represent more than 75 percent of reported cases between 1995 and 2010. Costa Rica and Nicaragua were included because of their relative high incidence. Finally, Colombia was selected from the third group.

2.2 Climatic Limiting Factors

Climate directly influences the likelihood of appearance of a vector-borne epidemic outbreak. Geographical and climatic conditions are, moreover, structural determinants of vector population evolution and survival rates. Table 2 describes the temperature limiting conditions for the survival of various vector-borne mosquito populations. The literature suggests that two major sets of variables influence the prevalence of dengue and malaria in a particular region.

2.2.1 Weather and Climatic Variables

Temperature, precipitation and humidity are the most relevant climatic variables, while weather involves the likelihood of short-term deviations from multiyear trends (Johansson, Dominici and Glass, 2009a and 2009b). Precipitation has a direct effect, as mosquito population can only thrive if rainfall is present. There seems to be a lag between above-mean periods of precipitation and reported outbreaks of the disease. Roseghini et al. (2011) and Johansson, Dominici and Glass (2009a) perform statistical analysis to identify the lengths of such lags; however, their results cannot be generalized, as these numbers are highly dependent on other variables. Given that temporally aggregated measures are better understood than daily ones, it is reasonable to use weekly, decadal or monthly precipitation measurements rather than daily records to capture extreme weather effects that may lead to vector-borne outbreaks.

Higher temperatures, for their part, reduce the time required for the pathogen (a virus in case of dengue, and a protozoan in the case of malaria) to replicate and disseminate in the mosquito (see Table 2). This process, referred to as the “extrinsic incubation period,” must occur before the virus can reach the mosquito’s salivary glands and be transmitted to humans. If the mosquito becomes infectious faster because temperatures are warmer, it has a greater chance of infecting a human before it dies (Dhiman, Pahwa and Aditya, 2008). Therefore, temperature may modify the growth of disease-carrying vectors by altering their biting rates, as well as affect vector population dynamics and alter the rate at which they come into contact with humans. Finally, a shift in temperature regime can alter the length of the transmission season (Patz et al., 2005).

2.2.2 Serotype

Mosquito serotype populations tend to be regionally clustered for a variety of reasons, although geographic and climatic factors are by far the most important determinants. In general, certain populations tend to create resistance to endogenous serotypes, particularly for dengue virus serotypes. However increased human mobility is changing this feature; when new serotypes are introduced into a region, the population becomes more susceptible to the new variety of mosquitoes (Roseghini et al., 2011). While relevant, this parameter cannot be included in the empirical model (Section 5) because records reporting the main serotypes causing outbreaks are not available in the sources consulted.

Climate change may shift the pattern distribution of vulnerability across regions in Latin America, as dengue and malaria require specific climatic conditions for growth (Githeko et al., 2000). Increased levels of rainfall and variations in temperature regimes across regions resulting from climate change may create conditions favorable to the appearance of susceptible mosquito populations. In fact, regions where those diseases have no historical presence may suffer disproportionately if appropriate policy measures are not implemented.

2.3 Literature Review: What Does Evidence Say about the Role of Climate in Health Outcomes and Vector-Borne Diseases?

The causal relationship between climatic extreme events and health outcomes has been analyzed mainly in epidemiological research and, more recently, in an emerging literature in economics. Some general conclusions can be drawn from this analysis.

Within the economics literature, there is a growing number of efforts to assess the effects of extreme events on health outcomes. Most of them relate extreme temperatures to some measure of mortality. Deschênes, Greenstone and Gurryan (2009) estimate the impacts of exposure to extreme hot temperatures during pregnancy, finding a strong negative relationship between birth weight and temperature. Deschênes and Moretti (2007) relate mortality measures to increases in spikes in temperature (and to extreme low temperatures). Among their results, they obtain a measure of locally based heterogeneity, finding that the effect of extreme events on mortality for counties in the bottom decile is 66 percent higher than those in the first decile); they also found sizeable differences among groups. Deschênes and Greenstone (2011) show the temperature-mortality response function estimated from daily temperature and annual county-level mortality data for the United States between 1968 and 2002. The evidence is generally suggestive of geographical differences in the effect of high temperatures on mortality, measured in different manners.

Extending the work of Deschênes and Greenstone (2011), Burgess et al. (2011) carry out perhaps the most comprehensive study of the impacts of climate on health outcomes in the developing world. The authors offer two clear causal mechanisms. First, extreme weather events have a direct impact on health via disease burden and psychological stress. Second, indirect effects include an income-based channel and a consumption-based channel; the first set of effects stems from the fact that suffering a disease will entail loss of working days with subsequent impact on household income-generating capabilities, while the consumption-based channel is the consequence of shortages in food provision as shocks are propagated to the rest of the economy, which may reduce the possibility for smoothing. Both phenomena may contribute to excess mortality.

Burgess et al. estimate district-by-age group models for both economic and health outcomes (mortality rates, wages, labor supply, etc.), estimating the effects under a dynamic optimization problem. They include bins and splines for both temperature and precipitation to allow nonlinear effects of climate on outcomes. The authors take advantage of extensive panel data at the district level (1957-2000) and include district area and year-by-year fixed effects, as well as region-specific time trends and daily data on temperature and precipitation. Guerrero (2013) replicates the methods and analysis used by Burgess et al. for all municipalities in Mexico for the period 1980-2010.

The evidence with respect to precipitation and humidity is scarcer. Barreca (2009), however, has used panel data for the United States for 35 years from over 350 counties. The author also provides evidence on how low humidity levels are strong predictors of deaths associated with influenza. The methods utilized in this paper are similar to those discussed by Deschênes and his coauthors.

A smaller set of papers within the economics literature use different methodologies to assess climatic impacts. Bosello, Roson and Tol (2005) analyze the economy-wide effects of climate change using a multi-country CGE model calibrated for 2050. Tol (2000a and 2000b) estimates the impacts of climate change under a cost-benefit framework.

In the epidemiology literature, the most common empirical approach is to model the intensity of outbreaks as a Poisson process and variants of this type of models. Here the literature is clearly more extensive, and there are a relatively large number of studies for vector-borne diseases as well. For the case of dengue, Johansson, Dominici and Glass (2009b), using 20 years of data and a statistical approach to control for seasonality, show a positive and statistically significant association between monthly changes in temperature and precipitation and monthly changes in dengue transmission in Puerto Rico. Similarly, Johansson, Dominici and Glass (2009a) analyzed the relationship between ENSO, local weather, and dengue incidence in Puerto Rico, Mexico, and Thailand using wavelet analysis to identify time and frequency-specific association. Their results indicate the importance of regional and local effects: even though dengue viruses have a universal transmission cycle, changes in temperature or rainfall may have diverse local effects. Roseghini et al. (2011) aim to analyze incidences of dengue fever in three different cities in Brazil. Monthly and seasonal timescales of dengue epidemics in relation to climate (temperature) were first analyzed by comparing time series of absolute values and anomalies between dengue and climate factors. Medina-Ramón and Schwartz (2007) focus exclusively on the direct health effects of changes in temperature, and particularly on the impact on mortality; they particularly note changes in the frequency of extreme weather events such as increasing extremely hot days and decreasing extremely cold days.

All the papers reviewed within the economics literature share some common features. First, they assess direct effects of climate extremes on health outcomes (mostly mortality). Second, they take advantage of large panel datasets to analyze within-country differences, which allows them to test for regional/geographic differences as well as for heterogeneous effects among subpopulations. Third, they generally analyze daily weather and account for nonlinear effects in their estimations. Fourth, stemming from the fact that climate is almost by definition an exogenous process, their results are internally and externally valid, at least for the United States and countries with similar socioeconomic and climatic conditions. The general conclusion seems to point to a non-negligible effect of extreme climate events on health and related outcomes; this seems to be well documented, at least for developed countries, with the notable exceptions of Burgess et al. (2011) and Guerrero (2013).

2.4 Literature Review: What Does Evidence Say about Social Conditions, Health Outcomes and Vector-Borne Diseases in the Context of Climate Change?

The literature linking urbanization, socioeconomic conditions and health outcomes is vast. The evidence found in the literature is robust with regard to this causal link. Nonetheless, some services will tend to have a larger impact on the likelihood of an outbreak than others; sewerage, access to piped water and provision of health services, for instance, are expected to be better predictors of outbreaks than other services (Kjellstrom et al., 2007; Vlahov et al., 2007).

Furthermore, differences within regional spaces definitely play a role in the impacts epidemic outbreaks can have on different population groups within a region (Galdo and Briceño, 2005). The availability of some public services varies considerably across cities, particularly in large and intermediate metropolitan areas, which constitute a vast majority of urban settlements in Latin America (IDB, 2011; Carrera, 2013). In this context, the most vulnerable groups will often have little or no assets easily convertible into money, as well as limited access to publicly provided services (Ruprah, 2009; WHO, 2010; Northridge and Freeman, 2011). So far, experience in the developing world has demonstrated that urbanization is accompanied by inequality and exclusion.

Demographic, behavioral and social factors are often keys for effective communicable disease control and underpin successful public health programs. Population features that are determinants of vector-borne diseases outbreaks include the density of a particular region, a precondition as it facilitates disease transmissions (Medina-Ramón and Schwartz, 2007). The increase in human settlements in cities creates scenarios that enhance the spread of the disease vector. For instance, Roseghini et al. (2011) observe that many cities do not provide adequate garbage collection, which results in construction debris and tires accumulating on properties.

However, these factors are not completely understood in the case of dengue (Guha-Sapir and Schimmer, 2005) and other vector-borne diseases (Hii, 2013).

Climate change will likely aggravate already existing urban social inequities and health risks, thereby exacerbating existing urban health inequities. Generally presumed impacts of climate on health in urban contexts are presented in IPCC (2007). The general conclusion is that certain population groups are at risk; children, pregnant women, and the elderly are generally more susceptible, especially for heat and weather-related illness and death, vector-borne and zoonotic diseases, and waterborne and foodborne illnesses. Some of the studies cited in the context of climate analyzed data differentiating among socioeconomic groups in urban contexts (Deschênes and Greenstone, 2011; Barreca, 2009). These results are expected to be more nuanced in cities in low and middle-income countries (Friel et al., 2011; WHO, 2010).

Despite this reality, the literature consulted suggests that further evidence at the community level is necessary to isolate climatic from socio-demographic determinants (Guha-Sapir and Schimmer, 2005; Kovats et al., 2000). As Friel et al. (2011: 886) stated: “There is still much to learn about the extent to which climate change affects urban health equity and what can be done effectively in different socio-political and socio-economic contexts to improve the health of urban dwelling humans and the environment.”

Several factors that are relevant to health outcomes, regarded as the social urban environment (Galea and Vlahov, 2005), remain out of our analysis. These include but are not limited to factors such as spatial segregation and inequality. While they are relevant, they are also difficult to measure at the scale presented here, and therefore, it is natural to believe that they will be part of the unobserved portion in our model. Thus unobserved heterogeneity presents a problem of omitted (and likely relevant) variables, which calls for taking action in this regard.

3. A Framework to Analyze the Impacts of Climate Change and Socioeconomic Factors on Vector-Borne Diseases

The discussion presented above highlights the importance of differentiating the effects of climate from those of socioeconomic factors on vector-borne disease outbreaks in particular and on health vulnerability in general. An alternative based on quasi-experimental methods of impact evaluation is presented here. Central to our estimation strategy is the notion that climatic variables are exogenous, as well as independent from socioeconomic covariates. In order to obtain measures of the effects our approach is divided into three steps.

The main question that this research aims to answer is *what is the effect of suffering an outbreak (of a given intensity), conditional on climatic and socioeconomic conditions, has on health vulnerability?* In order to achieve this goal, we take advantage of a commonly accepted notion of vulnerability and matching techniques frequently used to estimate impacts in observational studies.

3.1. Vulnerability

A commonly accepted definition of vulnerability is as the product of exposure and sensitivity which, along with adaptation capacity once a malaria or dengue outbreak takes place, affect the resilience of the community (Few, 2007; Miller, Yoon and Yu, 2013; WHO, 2003). Using this definition, a measure of vulnerability for the fourth administrative layer within a country is presented below.

Denote the four sets of political-administrative levels by J, Ω, Ψ and P, denoting the fourth, third, second and national layers, respectively. These sets are related to each other by $\mathrm{J} \subset \Omega \subset \Psi \subset \mathrm{P}$. If the analysis is performed within a single country, $P = 1$. Let g be one out of J fourth-level administrative units within region $\mathrm{G} \in \Omega$. Therefore, the local vulnerability measure for community g and period t can be defined by equation (1). As previously mentioned, this vulnerability measure will be dependent on socio economic and climatic factors; however, sensitivity will only depend on socioeconomic covariates, conditional on suffering and outbreak (of a given intensity).

$$Vulnerability_{gt} = \gamma(I_{Gt} | \, \mathrm{Clim_{Gt}}, Soc_{Gt}) \cdot Sens\_ind_{gt}\big(Soc_{gt} \big| I_{Gt}\big) \quad (1)$$

In equation (1), $\gamma_G(\cdot)$ represents the reduced form probability of an episodic outbreak in the aggregated region, G. The appropriate definition of regions depends on the availability of climatic and epidemic information. In general, both pieces of data are available at the city level. I_{Gt} can be either a bivariate or a multivariate process, representing only the occurrence of an outbreak or its intensity, respectively. For simplicity, the time dimension will be excluded hereafter unless explicitly stated.

More disaggregated local measures of the probability of an outbreak could even be constructed at the community level, g; however, the information requirements for doing so would substantially increase. The multitude of factor involved in the emergence of an outbreak condition. The alternative consists of developing a structural model for the process of diffusion of the disease within the population in *g*[4] following Martcheva and Prosper (20011), Derouich, Boutayeb and Twizell (2003), or Hethcote (2000). Risk maps such as MARA LITE for malaria and CIMSiM and DENSiM for dengue can additionally support the analysis of local outbreak case studies.

The second term, $Sens\_index_{g,}$ is a multidimensional index of sensitivity, which includes adaptation measures such as those proposed by OECD (2008) and implemented by Hagenlocher et al. (2013). It aims to capture in a single measure the multiple aspects that negatively influence the impact of an outbreak, and the strengths that help to halt the impact of epidemics within populations.

Vulnerability can be the result of a deficient provision of public goods and households' lack of capabilities to cope with adverse events. Therefore, the index should include variables that capture differences in personal and community assets. The former consist of households' social and economic assets, while the latter consist of both privately and publicly provided services: hospitals and other health facilities available nearby, access to basic services and network coverage, and quality measures of these services.

In this multidimensional framework, it is necessary to introduce adaptation measures. Sensitivity can be regarded as the opposite of resilience and adaptability. Therefore, the covariates included have a clear direction. This conceptual framework may open the opportunity to build vulnerability indexes based on local data gathered from quality of life indicators. Among other data sources, the IDB's Sustainable and Emerging Cities Initiative can enhance the efforts of local movements to create information that is comparable across cities.

Ideally, the index should be able to capture the effect of the intensity of an outbreak. A city or community may be resilient up to certain point but become completely vulnerable afterwards, depending on the intensity of the shock. This is particularly relevant for epidemic outbreaks, as local health facilities can collapse once certain thresholds are reached, and nonlinearities of this type should be included in the modeling and construction of the sensitivity index. However, the index presented here is not capable of capturing this relevant feature.

[4] For instance, this would require us to model outbreaks with contagion models, based on an ergodic Markov process, such as the SIR (Susceptible-Infected-Retired) and SIS (Susceptible-Infected-Susceptible) models.

Intuitively, expression (1) argues that populations within a region *G* will be exposed to the same risk of an outbreak, but that the local economic and social conditions will shape the magnitude of the impact in a particular community *g* within region *G*.

3.2. Establishing Causal Mechanisms

The conceptual framework for vulnerability discussed above provides the analytical conditions for testing two different mechanisms that would indirectly impact on health outcomes by increasing the vulnerability of communities. Figure 1 describes the proposed mechanisms. In all cases, a first order negative relationship between exposure and sensitivity to health outcomes is assumed (i.e., higher levels or vulnerability will translate into poorer health outcomes). The mechanisms considered are the following:

1. *Direct mechanisms affecting the likelihood of an outbreak.* Both climatic and socioeconomic factor are combined to create the conditions for an epidemic outbreak. The economics literature reviewed in Section 2 makes the case for extreme events affecting health outcomes; in the case of vector-borne diseases, this means that an epidemic outbreak must occur in order to observe health outcomes (affected population, morbidity and mortality). Similarly, urban conditions can have a significant impact on the likelihood of an outbreak. Poor, densely populated communities that lack access to sanitation are expected to experience epidemic outbreaks more frequently than better-equipped communities.
2. *Direct mechanisms impacting community sensitivity.* As mentioned, sensitivity is the result of city-wide, community and household-level capabilities that enables individuals and households to cope with adverse events. Following the ideas presented in Cutter, Boruff and Shirley (2003) and IPCC (2007), this framework allows us to estimate an index composed of two different dimensions (a household-level factor and a community/city factor) that add up to represent the abstract concept of sensitivity.
3. *Institutional and other confounding factors.* Regrettably, not all relevant variables can be included in the model, mainly because they are not necessarily observable. While the quality of institutions within a city, region

and/or country has profound effects on the quality of available health and infrastructure services, institutional quality is seldom measured at the city level (third administrative layer). Similarly, mosquitoes' serotypes, local microclimates and inhabitants' habits and culture will almost certainly affect the likelihood of an outbreak but cannot be observed, at least at the scale proposed in this paper.

Under the hypothesis established by these mechanisms, the objective will be to isolate the effects of exposure using a set of propensity score measurements in order to directly associate vulnerability with sensitivity. Once treated and non-treated regions with similar probabilities of suffering an epidemic outbreak of dengue and/or malaria are identified, the differences in their capabilities to cope with the adverse event will determine their vulnerability and therefore their expected health outcomes. In this regard, the outcome variable will be the sensitivity index. A more detailed exposure of these ideas is presented below.

Acknowledging the complex nature of episodic outbreaks, the crucial assumption in the approach proposed here is that local environmental conditions are statistically independent from local institutional capabilities for coping with climatic risk, along with the assumption that climatic factors are strictly exogenous to vulnerability (and therefore health) outcomes. Socioeconomic conditions, however, are allowed to be endogenous. For relatively short periods of time, spanning from years to some decades, these assumptions hold for the empirical evidence obtained.[5] Defining the errors of the probability model for an epidemic outbreak as ε_G, the assumptions can be stated as:

$$E\left[Soc_{Gt}, Clim_{gt}\right] = 0 \ , for\ all\ G, g\ and\ t \qquad (A.1)$$

$$E[\varepsilon_{Gt} |\ Clim_{Gt}] = 0 \ , for\ all\ G, g\ and\ t \qquad (A.2)$$

$$E[\varepsilon_{Gt} |\ Soc_{Gt}] \neq 0 \ , for\ any\ \ G, g\ and\ t \qquad (A.2)$$

$$\gamma(I_G |\ \mathrm{Clim}_G, Soc\ _G) = \delta \cdot Clim\ \ _G + \beta \cdot Soc\ _G + \varepsilon_G \ (2)$$

Let equation (2) be the reduced- form linear probability of an outbreak conditional on a set of climatic and socioeconomic covariates. The residual ε_{Gt} follows an i.i.d extreme value distribution. Therefore it is possible to obtain the vector of marginal effects of the set of

[5] This might not be the case for long-term (centuries) analysis. Acemoglu, Johnson and Robinson (2002 and 2003).

variables. For a given number of climatic covariates C, the marginal effect of a particular regressor $c \in C$ on the probability of an outbreak in region G, I_{Gt}, can be denoted by $\frac{\delta\gamma(\cdot)}{\delta Clim^{c}}$. Similarly, the marginal effect of a socioeconomic covariate, $s \in S$, is denoted by $\frac{\delta\gamma(\cdot)}{\delta Soc^{s}}$. This set of equations will be used to estimate the impact of certain climatic and socioeconomic covariates on the exposure component of vulnerability.

$$\frac{d\gamma(I_G|\,\mathrm{Clim_G}, Soc_G)}{dClim} = \begin{bmatrix} \frac{\delta\gamma(\cdot)}{\delta Clim^{c=1}} \\ \vdots \\ \frac{\delta\gamma(\cdot)}{\delta Clim^{c=C}} \end{bmatrix} \quad (2.a)$$

$$\frac{d\gamma(I_G|\,\mathrm{Clim_G}, Soc_G)}{dSoc} = \begin{bmatrix} \frac{\delta\gamma(\cdot)}{\delta Soc^{s=1}} \\ \vdots \\ \frac{\delta\gamma(\cdot)}{\delta Soc^{s=S}} \end{bmatrix} \quad (2.b)$$

Likewise, under assumption (A.1), the conditional probabilities in either the climatic or the socioeconomic dimensions can be defined by expressions (2.c) and (2.d), where $\tilde{x}$ represents the mean value for variable x. These will be called partial conditional probabilities and will provide a mechanism to isolate the effect of one set of covariates.

$$\gamma(I_G|\,\mathrm{Clim_G}) = \int \gamma(I_G|\,\mathrm{Clim_G}, Soc_G)dSoc = \gamma(I_G|\,\mathrm{Clim_G}, Soc_G = \widetilde{soc}) \quad (2.c)$$

$$\gamma(I_G|\,\mathrm{Soc_G}) = \int \gamma(I_G|\,\mathrm{Clim_G}, Soc_G)dClim = \gamma(I_G|\,\mathrm{Clim_G} = \widetilde{clim}, Soc_G) \quad (2.d)$$

Equations (2.c) and (2.d) will be utilized as the propensity scores (Heckman, Ichimura and Todd, 1998; Caliendo and Kopening, 2005; Heinrich, Maffioli and Vázquez, 2010; Hirano and Imbens, 2004) that will enable us to construct an artificial control and therefore estimate the treatment effects of both climatic and socioeconomic variables on the vulnerability measure defined by equation (1). Equations 2.c and 2.d will allow us to estimate the potential effects of marginal changes in climate conditions or improvements in socioeconomic indicators on the likelihood of an outbreak (see Figure 1).

3.2.1 Marginal Effects on Exposure (MEE)

These estimators can support the analysis of *changes* of certain indicators within the vector of either social or climatic covariates. It is based upon estimating propensity scores in (2.a) and (2.b). They will be utilized to analyze impact of a single variable on the exposure component of our vulnerability measure—for instance, to assess the effect of a 1 percent increase in the rate of access to sewerage service on the likelihood of an outbreak (of a given intensity). Odds ratios in the case of binary treatments and incidence rate ratios (IRRs) in the case of a continuous treatment will serve as the measures.

3.2.2 Partial Treatment Effect on Vulnerabilities (PTE VUL)

Performing the matching algorithms on equations (2.c) and (2.d), it is possible to isolate the effects that *levels* of climatic and socioeconomic covariates can have on the proposed measure of vulnerability. Remember that for every observation in the sample, partial propensity scores for climate, for instance, will be estimated at the mean values of the vector of socioeconomic covariates.

$$PTEVul_{clim} = M[Sens\_index_G] - M[\widetilde{Sens\_index}_{G`}] \quad (3.a)$$
$$PTE\ Vul_{soc} = M[Sens\_index_G] - M[\widetilde{Sens\_index}_{G`}] \quad (3.b)$$

Before continuing, it is important to define the conditional probability when a region actually suffered and epidemic outbreak (the treatment) as $\gamma(I_{Gt}|\mathrm{T}=1,\ \mathrm{Clim}_{\mathrm{Gt}}, Soc\ _{Gt})$; and the probability of suffering an outbreak when it did not happen during the analysis period as $\gamma(I_{G`}|\mathrm{T}=0,\ \mathrm{Clim}_{\mathrm{G`}}, Soc\ _{G`})$, the control. Similar expressions shall be defined for equations 2.c and 2.d. A particular region *G* where *T*=1 will be compared to a single region *G'*, or to some weighted average of regions with similar propensity scores (see Section 5), denoted by $\tilde{\gamma}(I_{G`}|\mathrm{T}=0,\ \mathrm{Clim}_{\mathrm{G}\prime}, Soc\ _{G\prime})$.

For the treatment effects measures proposed below, both mean and dispersion statistics will be considered, in order to take advantage of community-level data and obtain a more comprehensive notion of the effects social and climatic covariates can have on vulnerability. The expected value of the sensitivity index for all communities in region *G* will become the sensitivity of region *G*; that is $E\left(Sens\_ind_g\right) = Sens\_ind_G$. Likewise, the weighted average of all comparison regions is denoted by $\widetilde{Sens\_index}_{G`}$. In the case of variance, the scalars denoting the summary dispersion measure are $Var\left(Sens\_ind_g\right) = V(Sens\_ind_G)$; the weighted average of the control group is $V(\widetilde{Sens\_ind_G})$. Mean, median, variance and percentile ratios considered will be referred as $M[\widetilde{Sens\_ind_G}]$ and $M[Sens\_ind_G]$ for control and treatment groups, respectively. The extension to estimating the average treatment effects on the treated is straightforward. Given our interest in analyzing the effect of an outbreak on health vulnerability, this latter measure is better adapted to these objectives.

Two relevant cases for analysis will result. The most interesting for our application will enable us to estimate the extent to which socioeconomic conditions affect the vulnerability of a city (equation (3.a)). The partial propensity score calculated, evaluated at mean socioeconomic covariates and actual climate conditions, will actually control for climatic variables to the extent that it will match cities with similar climatic and geographic conditions. As a result, the remaining elements affecting the vulnerability of a city will only be socioeconomic. In the second case (equation (3.b)), communities that have the same partial social propensity score (evaluated at mean values of climate covariates) will differ in the underlying climatic factors considered.

In order to illustrate this idea, consider the following simplified example. A set of cities can only experience two types of socioeconomic conditions, *s* and *s´*, but they all share similar climatic environments (in terms of mean temperatures and the number of extreme events experienced, for instance). Under the partial propensity scores evaluated at mean socioeconomic conditions, these set of cities will have very similar probabilities of suffering an outbreak if and only if they experience similar climatic and geographic conditions. Therefore, matching upon 2.c will enable us to obtain valid comparison groups after controlling for climatic conditions, and as long as the balancing conditions are fulfilled after matching observations, differences in vulnerabilities between treated and comparison groups can be entirely attributed to socioeconomic conditions.

3.2.3 Differences in Differences to Control for the Unobserved Heterogeneity

The measures of the treatment effects above are valid only if the unconfoundedness or conditional independence assumption (Rosenbaum and Rubin, 1983) holds. That is to say that, given the propensity score, both observable and unobservable characteristics are balanced among treatment and non-treatment groups, and therefore the potential issue of endogeneity is eliminated. Unfortunately, this assumption cannot be tested.

Furthermore, it is likely that unobservable (or immeasurable) institutional characteristics within regions are time invariant for the time span under consideration. This could lead to potentially biased estimators. For instance, it is more likely that within some regions experience more frequent episodes of epidemic outbreaks not just because of their climatic and/or socioeconomic conditions, but because of cultural practices favor the appearance of mosquito populations (e.g., structures or debris that hold standing water), or political institutions preclude appropriate investments in basic public infrastructure.

Given that counterfactual levels for treated and non-treated groups can be different because of these unobserved factors, at least the time variation should be similar between groups. Difference-in-differences argues for a counterfactual scenario by assuming that, in the absence of treatment, the change in treated outcome would have been like the change in the outcome of control group. Thus unobservable differences are regarded as time-invariant across/within cities or countries.

Galiani, Gertler and Schargrodsky (2005) use a combination of a PSM and differences-in-differences estimators to measure the impact of privatization of water and sanitation services on child mortality in Argentina in the late 1990s. Galdo and Briceño (2005), based on Ecuador's census for years 1990 and 2001, estimate the impacts on motherhood mortality of five different impact estimators based on propensity scores, along with differences-in-differences to control for unobserved heterogeneity across cities.

3.2.4 Extending the Framework to Account for Continuous Treatments

The conceptual framework presented above can be utilized to estimate the impacts not only of an outbreak's occurrence but also of its intensity. It is well understood in the epidemiology literature that epidemic outbreaks can have large nonlinear impacts on the number of affected individuals. Furthermore, under the conceptual framework presented above, sensitivity can be seriously compromised when outbreaks of different intensities take place. Regional health facilities and individual assets may not suffice to restrain the negative impacts of a massive outbreak.

In order to incorporate this relevant feature, Generalized Propensity Scores (GPS) techniques developed by Imai and Van Dick (2004) and Hirano and Imbens (2004), along with generalized extreme value (GEV) distribution assumptions for the residuals of the score model, will be utilized. The most common GEV distributions will be multinomial logit for ordinal-categorized data, and the commonly used family of Poisson distributions (negative binomial and zero inflated Poisson) to estimate the probability of given number of reported cases that can take place during a vector-borne outbreak.

3.2.5 The Sensitivity Index

The framework presented above relies on the notion that sensitivity is a continuous measure, which captures the different capabilities of cities and their inhabitants to cope with climate adverse effects on vector-borne vulnerability. As mentioned, vulnerability to health risks is composed of both exposure to an outbreak and sensitivity to that exposure. Once we control for exposure, differences in levels of sensitivity will translate directly into differences in vulnerabilities to vector-borne diseases. Under a central assumption we made, this will in turn translate into health outcomes.

To account for sensitivity, a measure of a region's capabilities and assets available to cope with adverse epidemic events needs to be created. Vulnerability to health risks is composed of both exposure to environmental changes and sensitivity to that exposure. The former can be measured directly using environmental data. However, sensitivity can only be measured indirectly with auxiliary socioeconomic indicator variables such as income, education and household characteristics; for this purpose a socioeconomic index is created. The main idea is to reduce all the socioeconomic information into a single variable ranking which captures all non-redundant information and where dimensions can be important. There are two important reasons for this: i) the created index variable can be used as a single weighting factor in further analysis; and ii) the index should explain the maximum amount of variability in the dataset with the least amount of linear redundancy (multicollinearity).

Following the conceptual framework developed in Cutter, Boruff and Shirley (2003), OECD (2008) and implemented by Hagenlocher et al. (2013), we estimate a sensitivity index composed of two different dimensions (a household-level factor and a community/city factor) that add up to represent the abstract concept of sensitivity. The index builds upon a set of underlying socioeconomic and demographic indicators.

Determining which indicators to include is a central problem when building the sensitivity index. A priori it is natural to assume that certain socioeconomic factors will affect sensitivity to potential health risks. What these factors are, however, is difficult to assess.

Socioeconomic indicators simply measure certain components of these hypothetical factors. Furthermore, many of these indicators are mutually correlated and may even be measurements of the same factor.

In order to determine which indicators to include, several techniques and approaches exist for analyzing multivariate data groupings and correlations are available. The ultimate determinants in choice of method are structure of data and end-goal. Exploratory factory analysis techniques including graphical methods such as Cluster Analysis are used mostly with a descriptive end-goal. For the purpose of creating an index, quantitative multivariate techniques are required (Abdi, Williams and Valentin, 2013). The procedures considered in this paper include Principal Component Analysis, Correspondence Analysis, Generalized Principal Component Analysis, and Multiple Factorial Analysis. A possible alternative strategy not considered is the use of regression techniques not discussed in this paper.

The method of Principal Component Analysis generates a set of mutually orthogonal vectors (uncorrelated vectors) which can be interpreted as the underlying socioeconomic factors. These orthogonal vectors can be interpreted as the underlying socioeconomic factors. Furthermore, a measure of variability is attached to these components which can be used to determine which components need to be considered for the analysis and which can be excluded. The sensitivity index is then created using the set of normalized indicator variables multiplied by the factor loadings associated with each indicator and the highest-ranked principal components. These components can be chosen with cutoff criteria for the eigenvalues or with the aid of graphical tools (Cuadras, 2012), such as a screen plot to see the contribution to explained variability.

Formally, let X be an n-by-k matrix with k (normalized) indicator variables. Then X can be decomposed into: X=VEU, where V is an N-by-N matrix of normalized vectors, E is an N-by-k diagonal matrix with eigenvalues (variances associated with each component), and U is a k-by-k matrix composed of the principal components which can be used as the weighting factors for each indicator variable. The index can then be expressed as the sum:

$$Sen_{Index} = \sum_i Z \cdot U_i \, , \qquad such\ that\ i = 1,2, \dots p \leq k \quad (4)$$

In equation (4) Z is the normalized data matrix with k indicator variables, and U_p is the last principal component determined to be used. Multiplying by the reduced eigenvalue matrix to add weighting information for each component can also be considered. Both alternatives are compared in this paper. Another major challenge with PCA is the choice of indicator variables to use as well as the final number of components to keep for the index; a discussion on the methodology adopted to address these issues can be found in Section 5.

Other methods such as Factorial Analysis, alternative measures of variance, and non-linear multivariate methods can be a viable option in certain cases. A close relative to PCA is Multiple Correspondence Analysis (MCA), used to find underlying factors when dealing with multiple categorical variables. The technique is similar, since MCA finds components which maximize "inertia," which is analogous to optimizing explained variability. Indeed, MCA and PCA results are highly correlated (Abdi, Williams and Valentin, 2013). Nonetheless, MCA should be used when dealing with categorical data. Unfortunately, MCA cannot deal with categorical and quantitative data simultaneously, which is the case with the census data used in this paper. However, there are related methods that deal with this problem. The most flexible and general form of PCA is that of Generalized Principal Component Analysis (GPCA). GPCA does not make any assumptions about the general structure of the data or, more importantly, the covariance (correlation) matrix used to determine principal components (Rao et al., 2007). The basic idea behind GPCA is to estimate correlations for the data such that the variance-covariance matrix yields more reliable estimates for the principal components. Thus it is more appropriate to use GPCA with mixed data types, especially when dealing with binary variables which may be highly unbalanced (highly skewed), which is the case with many of the indicator variables used to create the index in this paper. GPCA is a fairly recent technique and statistical routines found to deal with GPCA are user-written routines submitted to statistical packages. In this paper we constrain our analysis to the mixed PCA discussed; however, further research is warranted for comparison in construction of the sensitivity index.

As discussed earlier, one of the advantages of PCA is that the underlying factors need not be known a priori. However, once the principal components are established, there is the temptation to attach a certain meaning to each component once it has been estimated. Nonetheless, it is important to remember that components are only a result of the given sample and may change drastically in other samples. A certain degree of caution should be used in the interpretation of the components as they relate to the theoretical factors. Specifically, while the principal components may be used as weighting tools, one should attribute a specific interpretation to the components with caution (OECD, 2008).

Hagenlocher et al. (2013) consider two main factors as measures of sensitivity: "susceptibility" (SUS) and "lack of resilience" (LoR). Since these cannot be measured directly, a set of 12 indicator variables is used as a measure of susceptibility, and 11 indicators are used as a measure of resilience. These include racial information variables, age categories, education levels, household characteristics, and proximity to hospitals, among others. By performing a PCA, the authors are able to create an index based on the PCA factor loadings using the first two principal components as representative of SUS and LoR, respectively. Using census data for Cali, Colombia in 2005, Hagenlocher et al. (2013) find resulting weighting factors similar to those drawn from an "expert-based" approach used in their final analysis.

This approach, however, has three problems in relation to our current research. First, one must decide a priori which variables correspond to which factor—SUS or LoR. While this may seem intuitive, there is no reason to assume that certain variables belong to a single component, both components, or neither component. This is problematic because it both limits the usability of existing data and poses a greater challenge in obtaining data of limited availability. Instead, Principal Component Analysis can be used directly to determine how variables contribute to each underlying factor. Second, simple PCA requires that all data be quantitative (continuous) in nature. When dealing with categorical data such as that found in raw census data, alternative (albeit similar) methods must be employed to deal with underlying structure of the data. Finally, we do not have access to expert weighting factors. Instead, we must rely on a comparison of different statistical methods.

Vincent and Sutherland (2013) consider a similar approach but follow a different methodology for data reduction. In their initial step, they consider using an entire set of 79 variables to estimate the principal components for a socioeconomic status index for British Columbia, Canada. After finding implausible results, however, they follow a data reduction process by discarding highly correlated variables and count variables. They subsequently reduce the number of variables to 43 and use this set to perform a simple PCA. Unlike Hagenlocher et al. (2013), Vincent and Sutherland (2013) choose to create two different sets of rankings: one based on the two largest (in explained variability) eigenvalues (two sets of factor loadings) and one based only on the first principal component. Their reasoning is that, in using two principal components, they may be canceling out the effects of certain variables that appear twice with different signs or introduce a double counted weight with variables that appear twice with relatively large factor scores with the same signs. These are legitimate concerns and pose a challenge to creating such an index using the PCA approach. We consider similar approaches in this paper as well as principal component rotation techniques such as "varimax" rotation.

At the end, what is obtained from this process is a ranking, where distance among units of analysis is relevant but needs to be carefully analyzed. Which factors determine the position of a particular city in the ranking are therefore relevant for policy analysis. However, the quality of policy implication will heavily depend on the robustness of the results obtained for the sensitivity index under several calculation methods.

4. Data

In order to implement the framework introduced in Section 3, five countries were analyzed based on the availability of national census information disaggregated at the fourth administrative layer, and the number of cases and prevalence of cases reported by the PAHO for dengue fever (see Table 1).

Only dengue fever and malaria were selected for this analysis. As mentioned, dengue fever is the most prevalent vector-borne disease in the region, while malaria is a public health priority because of the large burden imposed on the infected population. However, the empirical analysis was only performed considering dengue information because of the scarcity of malaria data and dengue-carrying mosquitoes' largely urban habitat.

At some point socioeconomic information at the desired geographical disaggregation became the binding restriction on performing the analysis for all the countries considered in Table 1. Therefore three countries were prioritized based on the availability of socioeconomic information and their disease statistics reported. Brazil was included because of the large number of cases reported, and Costa Rica and Nicaragua due to their relatively high incidence of dengue fever incidence among its population. Mexico and Colombia, which display neither of these characteristics, were selected as the control group. Three different sources of information were consulted to extract the data for these countries. Table 3 summarizes the nature of the variables obtained and the sources from where these were retrieved.

4.1. Health Data

Health data were retrieved from the web portal Health Maps,[6] a web portal developed by a team of researchers, epidemiologists and software developers at Boston Children's Hospital, aimed at utilizing online non-official sources for disease outbreak monitoring and real-time surveillance of emerging public health threats.

Data for Mexico, Costa Rica, Brazil, Dominican Republic, Colombia, El Salvador and Nicaragua were manually retrieved for years spanning from 2006 to 2013. Tables 4 and 5 summarize the data obtained, which call for several comments. First, there is the reasonable presumption that Health Maps has an increasing reach and that more cases are being reported as time goes on. More than 70 percent of the reported outbreaks in the three countries considered, for example, took place in 2012 and 2013. This feature has important implications for the estimation strategy; it precludes analysis with panel data, as the basic conditions for balancing the distribution of observations cannot be fulfilled. In this case, the model described in Section 3 will only consider the probability of at least one epidemic outbreak. Health Maps presents a second important caveat which may endanger the validity of the results: given that only reported cases in news and/or official data are present in their databases, regions with inadequate media coverage or government records will be underreported in the dataset.

A second relevant characteristic of health data is the seasonal distribution of outbreaks (see Table 6). As was described in Section 2, dengue fever and malaria mosquito populations are heavily reliant on precipitation, temperature and humidity for their development and infectious behavior. The most favorable conditions for mosquitoes arise during the rainy summer season in the Northern Hemisphere. For Brazil, with a non-negligible share of its geographic extension in the Southern Hemisphere, the number of reported outbreaks is more evenly distributed.

Although serotype was reported as a relevant determinant of epidemic outbreaks, this parameter cannot be included in the empirical model since records reporting the main serotype causing outbreaks are not available in the sources consulted.

4.2. Climatic Geo-Referenced Data

The literature reviewed in Section 2 established the importance of geographic and climatic variables. Structural geographic variables such as the altitude, latitude and longitude, as well as

[6] http://healthmap.org

climatic distributions for precipitation, temperature and humidity, determine to a large extent the favorability of conditions for mosquito populations.

Two sources were used for monthly data for climatic variables from January 1998 to July 2013. Precipitation data were retrieved from NASA's Tropical Rainfall Measuring Mission (TRIMM) Satellite Databases, and mean monthly temperatures were obtained from the IRI climate data repository of Columbia University.

Several other papers have dealt with daily climatic data instead of monthly data (Burgess et al., 2011; Guerrero, 2013; Deschênes and Greenstone, 2011). Empirically, there are potential benefits and costs of using more aggregated data. On the downside, extreme events taking place over the course of a couple of days can be concealed by average monthly figures. On the positive side, the analysis with monthly data avoids tuning the model in order to capture the appropriate lags between climatic events and epidemic outbreaks, which can be very difficult to determine in a cross-country study given the complex local relationships that intervene (Roseghini et al., 2011); in this regard, monthly figures might capture more of the structural climatic conditions that determine epidemic outbreaks. In any case, the results obtained can be regarded as lower-bound effects, and extending the analysis to include daily data might constitute an important source of consistency and robustness.

Geo-referenced databases containing polygons down to the third political-administrative layer for every country in Table 1 were sought. GIS software was later used to assign each pixel of satellite data to those polygons. Three noteworthy cases resulted from this exercise, which are represented in Figure 2. In case 1, no information can be attributed to the polygon, and therefore no climatic information is available unless more complex interpolation techniques such as kriging are utilized; these observations were excluded from the analysis. Case 2 is the opposite, where several climate satellite point measurements fit within a political-administrative geographic region. In this case some sort of averaging technique should be used, and at this stage of the project the simple average was selected. Finally, the one-to-one case is the most useful, as a single satellite measurement can be assigned to each administrative level.

It is worth mentioning that of currently available climatic information for the original 12 countries considered in Table 1 exceeds 19 thousand points, as detailed by country in Table 7. However the final number of available climatic information is smaller because only a fraction of these pixels lie within case 3 in Figure 2. Aside from locations where no climatic information is available, where more than one point of climatic measurement lay within a location's geographic boundaries, simple averages were taken. The latter approach is not the optimal since it does not account for topographic and geographic aspects, which may become relevant in some circumstances where areas are sufficiently large to include different microclimates. The result of the procedure adopted is that, for instance, only 56 of initially available observations were finally available. Increasing the accuracy of the extrapolations by introducing kriging or other algorithms can be an important extension for future research.

4.3. Socioeconomic Data

Socioeconomic data were obtained from the national statistical bureaus of the five countries included in the sample, ECLAC's RADETAM platform and the Integrated Public Use Microdata Series (IPUMS). Information at the third and fourth layers of political-administrative division was collected for every country. Table 8 describes the sources, and Table 9 provides the list of outcomes included in the sample.

The selection of socioeconomic variables was in part determined by the availability of comparable information across countries. Although countries tend to ask the same questions in their national censuses, not all of the variables discussed in Section 2.2 could be found at the third administrative level in all countries. In this regard, proxy variables were included to capture similar effects. Where possible, correlation analyses were performed in order to validate the inclusion of those proxies. Summary statistics for the variables included in the sample, by treatment category, can be found in Table 9.

Income-generating capacity. Unemployment and illiteracy rate at the municipal level were included in order to capture the effect of poverty, given the lack of income or expenditure data that would have enabled us to construct a poverty index measure.

Access to and quality of sanitation and water services. Sewage access coverage (either to the public system or to a septic tank) and water supply (either within the dwelling or close to it) were included. Electricity coverage was included as well, despite the fact that this is the service with the highest coverage rates in the region (Mejía, 2013) and may have little impact on the appearance of vector-borne disease.

Vulnerable population and density. Direct measures for these variables are part of the sample. One is the share of population below 14 years old, as both dengue and malaria tend to have the largest impacts on this group (Patz et al., 2005; Ebi, Kovats and Menne, 2006). Density, measured in inhabitants per square kilometer, aims to capture the overcrowding effect that enhances virus transmission. There is, however, a caveat with this measure, as it was estimated using the political-administrative surface area, which is different than urban-populated area; this can bias the estimate, particularly in Mexico and Brazil, where huge municipalities are present along with relatively small ones.

Geographic controls. In order to control for the location of a region, both of a municipality's centroid coordinates (latitude and longitude) are included.

In addition to the socioeconomic covariates obtained from national censuses, country-level information from the World Health Organization, was incorporated in order to control for country effects, which are presumably relevant. The set of variables can be found in Table 10.

Using the health, climate and social datasets collected, simple and generalized propensity score matching were performed, and later differences in vulnerabilities were assessed based on a sensitivity index. The list of variables included was guided by the literature review presented in Section 2.

5. Results

5.1. Simple and Generalized Propensity Scores Matching

As described in Section 3, the estimation of causal effects from climate and social conditions to health vulnerability, and therefore on health outcomes, relies on first isolating the effect of the likelihood of an outbreak for a given city. This is achieved with two propensity scores proposed. One is the traditional simple propensity score (SPS), as proposed by Heckman, Ichimura and Todd (1998), where treatment can only take binary values. The second is a Generalized Propensity Score (GPS), based on the work of Imai and Van Dick and Imai (2004) and Hirano and Imbens (2004). The latter allows us to match locations with similar probabilities of experiencing an outbreak of a given intensity.

The unconfoundedness assumption (Rosenbaum and Rubin, 1983) is the cornerstone on which the propensity score technique is built. That is to say that, given the propensity score, both observable and unobservable characteristics are balanced among treatment and non-treatment groups, and therefore the potential issue of endogeneity is eliminated. Although this assumption cannot be directly tested, the so-called balancing conditions on observables between treated and non-treated regions help to validate the assumption. Several methods are proposed in the literature. Imbens and Rubin (2007) suggest testing for the normalized differences by the following expression: $(x_{g1} - x_{g0}) \big/ (s_{g1} + s_{g0})^{0.5}$, where x_{gT} are the sample means and s_{gT} represent the sample standard deviations for T= 0,1. The authors suggest the values for the standardized differences larger than 0.25 raise concerns that the treatment and control groups are not similar. Table 9 presents the results by country for the entire sample (pre-matching) by country. Using the Imbens and Rubin criterion, the results suggest that for the sample considered the pre-treatment differences in socio-demographic variables are considerable, suggesting that unobservable heterogeneity, presumably as the result of omitted variables, is present in the sample. This fact calls for taking care for unobserved heterogeneity using a differences-in-differences estimator (see Section 3). For climatic and geographic variables, the problem seems to be less pronounced.

We estimate a structural model for the probability of an epidemic (dengue) outbreak. Malaria was excluded from the estimation as the inherent process through which mosquitoes populations thrive are markedly different and outbreaks tend to be commonplace in rural areas. With this in mind, and considering the restrictions that health and climate data impose, which preclude a panel data analysis, all the measures for extreme climatic events described below were obtained for the entire sample of monthly weather data and later summarized for each and every location. The result is a dataset that allows us to estimate the probability that a particular municipality suffered *at least one epidemic outbreak* (SPS), or *an epidemic outbreak of a given intensity* (GPS) in the period 2006 through July 2013. In this regard, the models aim to capture some of the *structural* conditions necessary for epidemic outbreaks to take place.

Several extreme event measures were tested. Previous literature suggests that climate extremes tend to have nonlinear effects over health outcomes (Curreiro et al., 2002; Deschênes and Moretti, 2007; Burgess et al., 2011; Guerrero 2013); all the alternatives presented below aim to account for this relevant feature. For each of the alternatives, a different propensity score model was estimated. Only months from April to October were considered for Mexico, Costa Rica, Nicaragua and Colombia; for Brazil the distribution of climate variables for nine months in the year was included. This is consistent with the distribution of cases reported in these countries (Table 6).

*Model 1***.** First, following much of the work in the topic (Medina-Ramón and Schwartz, 2007), a dummy variable is used in case the monthly temperature average was greater than the 90th and 95th percentiles for each measurement location. An additional 75th percentile was included in order to capture the effect of “not so extreme” events. The final covariates under this model are the sum of these dummy variables for each location, or the number of months that experienced temperatures above the mentioned distribution thresholds.

Model 2. A second set of observations was included following the work of Burgess et al (2011) and Guerrero (2013). Ten bins for every 2 degrees from 10º Celsius onward were included. Similarly, 19 10-millimeter wide bins were included to account the effect of monthly accumulated rainfall, starting from 10 millimeters. As a result, the total numbers of months that lie within each bin are the covariates included under Model 2.

Model 3. A third characterization of extreme events was created using a cutoff for average monthly temperatures below 20 degrees Celsius, letting the values continue upward without restrictions (Medina-Ramón and Schwartz, 2007). In this case, both dummies for thresholds and continuous measures of climate (and their powers) above the thresholds were also included. Again, covariates considered under this model are the total number of months that fulfill the condition along with raw data for temperature and precipitation.

5.1.1. Simple Propensity Score

Results from the simple propensity score models are presented in Table 11a. The final models were selected based on the traditional Akaike (AIC) and Bayesian (BIC) information criteria; the BIC was prioritized given its large sample consistency properties, which are desirable when working with propensity score models. It is noteworthy that, since we are using logit and negative binomial models to estimate the probability of occurrence, nonlinearities are already present in the formulation. Additionally, standard errors were adjusted for intra-state correlations by using cluster-robust estimations at the second administrative level (the state or department level). A summary of the balancing conditions for the original propensity score is presented in Table 12; detailed results can be found in Appendix 1 (available from the corresponding author on request).

Several relevant conclusions can be extracted from Table 11, where results in bullets represent variable coefficients that are statistically significant at least at the 10 percent level.

First, there is a consistent impact of climate variables across model specifications. Physical and distribution thresholds considered for both temperature and precipitation are significant, consistent with results of Medina-Ramón and Schwartz (2007). The cumulative numbers of months above the 90^{th} and 95^{th} percentiles all have significant positive effects on the probability of an outbreak. Similarly, the cumulative numbers of months with average temperatures above 20 and 25 degrees Celsius, as well as cumulative number of months with average rainfall above 100 and 150 millimeters, all have significant impact on the likelihood of experiencing an outbreak. However, the effects turn out to be negative when average monthly temperature exceeds 150 millimeters of rainfall.

When it comes to temperature and precipitation bins, the results are not as straightforward as expected. Temperature bins are hardly significant, even though average yearly and summer temperatures are both statistically significant; their effects are similar in magnitude though opposite in direction, reflecting to some extent the physical-biological thresholds discussed above. In the case of precipitation, results are significant across the precipitation distribution (10 to 30 millimeters, 90 millimeters and 160 millimeters). Other structural climatic and geographic characteristics such as yearly and summer mean precipitation are also significant across models, suggesting what is long known, these are necessary conditions for outbreaks to take place.

Nonlinearities are also captured by the models, also consistent with some of the literature discussed above. However, the interpretation of such nonlinear effects in the context of an already nonlinear probability model is not straightforward. In model 1, the continuous terms after the cutoff points were all significant; these include terms in levels and their powers. In model 2, the significance of bins across the distribution adds up to account for these nonlinear effects on probability. The marginal effects exposed below can help to better explain such nonlinear effects.

For the case of socioeconomic covariates, the results are also consistent across model specifications. Total population, water supply and health expenditure per capita were all significant across specifications. The signs are theoretically consistent. It is noteworthy that density is only significant when total population is excluded as a regressor. In all of the specifications tested, the inclusion of these variables in logs outperformed their counterparts in levels.

The overall quality of the models considered is fairly good. The accuracy of the predictions (measured by the number of true positives and negatives) exceeds 90 percent under all models. Similarly, pseudo R-squared ranges from 0.28 to 0.37.

5.1.2. Generalized Propensity Score

In order to construct the Generalized Propensity Score, a negative binomial (Cameron and Trivedi, 1998; Hilbe, 2011) was used for the case of injured cases reported in Health Maps. The decision to use this parameterization was based on the analysis of the underlying processes of the variables to be explained. The Poisson distribution imposes an equidispersion condition that was not fulfilled in the data; the negative binomial regression estimates and additional overdispersion parameter allows relaxing the equidispersion assumption under a Gamma distribution representation.

The three different specifications that incorporate climatic effects were included. Results can be found in Table 14. The first notable aspect of the new results is that they are more sensitive to extreme events, reflected in the values of the parameters associated with the dummies for the 90^{th} and 95^{th} percentiles. Most of the patterns observed for bins are similar to those observed in the bivariate case, with the exception of temperature, where most of the bins across the temperature distribution are significant. However, the overall prediction power of any of the three models is relatively poor, at least compared to models under SPS, signaling perhaps that modeling more complex processes requires more detailed analysis and data.

5.2. Marginal Effects on Exposure

In the context of climate change, obtaining the marginal effects of climatic variables that capture the effect of extreme events, expected to increase their occurrence in the coming years, becomes relevant for policy analysis. Given the nonlinear nature of the model considered, several marginal effects were considered in order to obtain a more comprehensive notion of these effects on the likelihood of an outbreak of a given intensity. Table 13 presents the odds ratios (ORs) and average marginal effects (AMEs) for selected variables. As mentioned above, climatic variables were introduced in levels and socio-demographic data were introduced in logs in all the models for both SPS and GPS. In this regard, we are interested in analyzing marginal effects and semielasticities of the covariates described below:

- Climatic (marginal effects and semielasticities): i) number of months above climatic-physical thresholds under Model 1, ii) mean values for temperature and precipitation above physical thresholds of Model 2, and iii) number of months above distribution thresholds under Model 3;
- Socioeconomic (semielasticities): iv) total population, v) density, vi) health expenditure, vii) share of population below 14 years, and viii) water supply.

Our analysis concentrates on the average marginal effects, as they provide more information than marginal effects evaluated at any point (Wooldridge, 2010). Furthermore, as Wooldridge suggests, AMEs have a useful interpretation for policy analysis, as it is possible to retrieve the probabilities under both factual and counterfactual scenarios.

Climatic extreme conditions have a large impact on the likelihood of an epidemic outbreak. A 1 percent increase in the number of months registering temperatures and precipitations above the 90^{th} and 95^{th} percentiles has a significant impact on the probability of an outbreak of as much as 69 and 37 percent, respectively. If the average temperatures noted above were to increase by 1° C in the years ahead, cities in Latin America would see increases of as much as 2-3 percent in the likelihood of dengue epidemic outbreaks, assuming no future climate adaptation. However, and perhaps due to biological conditions, experiencing average temperatures above 25° C has a negative and large effect on the likelihood of an outbreak.

With regard to socioeconomic conditions, total population is by far the variable with the largest impact on probability, along with density. A 10 percent increase in total population and in population density is associated with a 66 percent increase in the likelihood of an outbreak. Similarly, any additional 1 percent of population lacking water supply services increases the likelihood of an outbreak by about 10 percent.

5.3. The Sensitivity Index

To account for sensitivity, a measure of a region's capabilities and assets available to cope with adverse epidemic events needs to be created. Following the conceptual framework developed in Cutter, Boruff and Shirley (2003), OECD (2008), and and implemented by Hagenlocher et al. (2013), we estimate a sensitivity index composed of two different dimensions (a household-level factor and a community/city factor) that add up to represent the abstract concept of sensitivity. The index builds on a set of underlying socioeconomic and demographic indicators.

Our sensitivity index was created by performing a mixed PCA on eight separate data tables made up of five Latin American countries at different census time periods. In our final analysis, we choose to keep the largest two components to create the index. The choice was based on the following criteria: i) consistency with previous literature, ii) consistent and sensitive results with the index, and iii) a high percentage of explained variances using the first two components. We compare simple sign-changing rotations for the principal components as well as varimax rotations, which coincide to a large degree.

Indicators for all countries include employment rate, average years of schooling, access to a sewage system, population under 14 years of age, and total population. Each country has additional indicators which we include in one index for comparison with an index created using only variables that countries have in common for all years

The results are fairly consistent across different time periods and countries, which is reassuring as a justification for using PCA to create the index. Figure 3 presents the results for Mexico. As can be observed, the most sensitive municipalities tend to be grouped in states where socioeconomic conditions are known to be the poorest within Mexico; the same happens with more resilient localities.

Table 15 presents the factor loadings for Brazil in the year 2000, which serves as an example of the final weights used to create the index for Brazil in that year (similar tables for the rest of the countries can be found in Appendix 2, available from the corresponding author on request). Summing the two components yields sensitive weights. The final value of the index is estimated using the loading factors weighted by the eigenvalues obtained by the PCA. Employment, education, type of worker, and ownership of household assets, such as televisions and refrigerators, make the greatest contribution to a *higher* ranking. Gender (female), number of children, disabilities, and lack of access to trash disposal services make the greatest contribution to a *lower* ranking. These results are similar to those found in previous literature, such as Vincent and Sutherland (2013) and Hagenlocher et al. (2013).

An issue for potential debate is how to define cutoff points and the relevant factors to be included in the final construction of the composite index. For instance, why unpaid workers earn a higher ranking than wage/salary workers can be questioned, and this is a legitimate concern. There are several approaches for resolving this issue. Variables that receive counterintuitive factor loadings could be removed, and a second PCA could be performed with the remaining variables. Another alternative is to apply different weights to each component using the respective eigenvalues for each component. That would give a higher contribution to ranking to higher number of years of school as opposed to unpaid workers, but it would still leave unpaid workers with a higher ranking than wage and salary workers, since both display the greatest value in the second component. Multicollinearity might also be influencing the results, but as the correlation matrix in Appendix 2 shows there is little need for concern since the highest correlation amongst variables is below 0.75, and the majority of correlation coefficients are below 0.30 in absolute value. For these reasons, and in an effort to account for the highest amount of variability, we decided to include all variables in the final analysis.

A better understanding is gained by observing the biplot for the factor loadings for Brazil 2000 and Mexico 2000 included in Figure 4. The plot is a visual representation of the factor loadings. The axes are the first and second principal components (SUS and LoR). The plot clearly shows how variables are projected onto the principal components. The first principal component is positively projected onto by number of bathrooms in the household, average years of schooling, ownership of a refrigerator and ownership of a telephone. No trash service, being self-employed, number of children and disabilities project negatively onto the first component. The first principal component explained roughly 16 percent (eigenvalue of 5.01) of the total variation in the dataset. We interpret this component as the measure "infrastructure" or, to be consistent with previous literature, as a measure of "susceptibility."

The second principal component is composed of employment status, worker type, and child survival rate in the positive direction, and number of children and female gender in the negative direction. We interpret this second component as a measure of "condition" or "resilience," to be consistent with the literature. Electricity is an interesting feature in the interpretation of the factor loadings. Despite being the fourth highest in contribution to the second principal component in the negative direction, the plot clearly shows that it is actually projects more onto the *first* principal component, as we would expect, since electricity is more a measure of infrastructure than condition (although it could be correlated with both). This visualization makes the variable projections easily classifiable into the two principal components. The table following the biplot summarizes this as well by sorting the factor loadings for each component.

5.4. Treatment Effects on Vulnerability

Based on the parameters estimated with the three SPS models discussed, propensity scores were as defined by equations (2), (2c) and (2d). The main statistics for each of these scores are presented in Tables 11 and 14. It should be mentioned that those scores will be assumed to be known before treatment, although in the case of socioeconomic conditions they can only be observed after the treatment actually took place.

As mentioned before, the unconfoundedness assumption is central to propensity score matching: by means of the propensity scores all observed differences between treated and control groups are eliminated, and so unobserved differences should be removed as well. Thus the starting point in order to assess the validity of the evidence found is to analyze whether balancing conditions are fulfilled. Two alternative matching algorithms were used for each of the three different propensity scores constructed. Balancing tests were then applied. A detailed description of the balancing tests for every variable and model considered can be found in Appendix 1 (available on request from the corresponding author). The original propensity scores under accomplish balancing conditions for both types of covariates, except for some socioeconomic variables in Model 3. In this regard, the average treatments on the treated effects reported in Table 16 are valid measures of the effects an epidemic outbreak can have on sensitivity. In the case of partial climate propensity, the balancing conditions are naturally accomplished for the case of climatic covariates. It is important to remember that the partial propensity scores are built in order to isolate the effects of either social or climatic conditions. In the case of the climatic partial propensity score, the predicted probability for each location is estimated using the mean values of socioeconomic covariates and letting the climatic variables vary across locations; in this regard matched observations will have comparable climatic factors in order to have similar scores. However, balancing conditions fail for some relevant socioeconomic variables under partial climate propensities. The situation is markedly different for the case of the partial social propensity. In this case, neither of the matching algorithms could ameliorate before-treatment differences—not even for socioeconomic variables. The latter two cases call for an additional step in order to obtain consistent estimates of the differences in health vulnerability across locations; estimation of average treatment effects via differences-in-differences is used to overcome this issue.

We are interested in those effects reflected by the average treatment on the treated as a measure. Once the effect on health vulnerability has been isolated via the propensity scores, the remaining differences in vulnerability can be attributed to differences in the sensitivity measure, as explained above. Thus results represent the impact of an outbreak on a city's health vulnerability. Treatment effects are reported for the case of a 10-neighbor, 10 percent caliper, matching algorithm as in Leuven and Sianesi (2003); a summary of those results can be found in Table 16. The application of kernel matching algorithms yield similar results, but they are not reported for the sake of brevity.

The results presented in Table 16 consistently signal that there are no significant differences in sensitivity between units that experienced an outbreak and those that did not during the sample years, when a valid score is used to match observations. Combining the exposure measures along with the constructed sensitivity index discussed above, it is possible to obtain a municipality-based measure of vulnerability to vector-borne diseases. A very preliminary analysis of the main features associated with vulnerability is performed by dividing the sample into the 30 most vulnerable municipalities, the 30 most resilient, and the rest intermediate. Results reported also in Table 18 suggest that the most significant differences between stem from differences in the propensity scores, reinforcing the importance of the marginal effects discussed above.

As mentioned in Section 3 and verified by the balancing conditions above, the conditional independence assumption does not hold when considering socioeconomic covariates. In order to deal with this sort of unobserved heterogeneity, which may induce endogeneity, a difference-in-differences estimator is proposed. The differences-in-differences coefficient was estimated using the sensitivity indexes presented above for census years 2000 and 2010 for Brazil, Mexico and Costa Rica, the only countries for which two census waves are available. The conventional common trends assumption for treated and control groups is tested in Table 17a and 17b; the results satisfy the notion that the evolution in covariates was similar both groups during the intra-census period.

The results suggest that, in terms of the health vulnerability to dengue measure proposed in this paper, there are no remarkable differences between treated and non-treated locations. The outcomes, presented in Tables 16, 17a and 17b, were tested for both the mean and deviation of the sensitivity index, using the three different scores discussed above as matching devices; the generalized propensity score was excluded from the analysis. Once the effect of exposure to an outbreak is removed, then there are no statistically significant differences between communities that suffered an outbreak and those that did not. Considering that the sensitivity index is built only from socioeconomic data, it is reasonable to obtain such a result. This fact raises the importance of building a sensitivity measure that reflects the magnitude of the impact. A function linking the observed magnitude to the local strengths of a community is necessary in order to improve the results. This is not to say, however, that the results are not valid for analyzing the underlying factors determining cities' capabilities for coping with adverse effects.

6. Conclusions and Policy Implications

The paper demonstrated that both climatic and socioeconomic conditions can have non-negligible effects on health vulnerability to vector-borne diseases. We achieved this by taking advantage of a commonly accepted notion of vulnerability, made up of two channels: a probability that accounts for exposure to an epidemic outbreak, and a sensitivity term intended to capture a city's capabilities for dealing with adverse events. As a result, the paper provides evidence that structural climate and socioeconomic conditions play a significant role in the appearance of epidemic outbreaks. The paper additionally identifies the most relevant socioeconomic conditions, all amenable to public policy, which make up the sensitivity measure used in this paper. Several policy conclusions can be drawn from this evidence, and possible avenues for further research are discussed below.

Climate extreme events do have a non-negligible effect on the likelihood of an epidemic outbreak of dengue. The results are consistent across three types of models introducing different climatic conditions. All these measures represent the effects of extreme weather events on the likelihood of a dengue epidemic. In the context of climate change, these numbers can help policymakers in their quest to quantify effects on human health in light of the extensive literature assessing the impacts of epidemics on health outcomes (Molina, 2009; Tarazona and Gallegos, 2011) and therefore guide policy initiatives.

In the case of socioeconomic conditions, the effects are also significant. As expected, water supply—one of the basic services that presents the greatest challenges for urbanization—is a relevant determinant of epidemic outbreaks. Expanding access to water supply services by 1 percent can reduce the probability of a dengue outbreak by 10 percent. Other socioeconomic variables, such as population (and density) and average per capita expenditure on health services are also relevant.

A sensitivity index, composed of two dimensions, was constructed upon socioeconomic variables at the household and community level, obtained from national censuses for five different countries. Our results are consistent with previous literature on the topic, and the ordering of cities based on the index is consistent with the observed distribution of socioeconomic indicators in selected countries, which validates our results. Nonetheless, further analysis is necessary, under different assumptions and models, in order to check for the robustness of the results obtained.

The validity of the evidence found is additionally noteworthy. Given the large number of cities and communities considered (more than 1,800) across five countries where dengue is a national health priority, the results enjoy a remarkable degree of external validity, at least for Latin America and the Caribbean. Similarly, the internal validity of the results is guaranteed by the empirical strategy followed; the propensity scores utilized achieve balancing conditions and, where these are not accomplished, differences-in-differences contribute to validating the results.

The framework presented can be expanded in several directions. In order to obtain a more local measure of exposure, structural modeling of the disease process could be obtained. In this paper, climatic data from satellite sources was processed using simple averages by city. In order to gain precision on the effects of climate variables, more geographically sophisticated models of climatic data interpolation are necessary, which should also be able to capture extreme events, though. Furthermore, regional climate models can be incorporated to analyze the long-term effects of climate change. Other types of vector-borne diseases and respiratory outbreaks can likewise be analyzed, providing a more comprehensive assessment of the impacts of climate change on health.

Finally, knowledge and data on both components of the vulnerability measure utilized has been acquired, which has enhanced our understanding of the problem. We are now in a position to analyze more carefully the distribution of vulnerability in order to integrate the analysis into obtaining policy relevant conclusions. For instance, it would valuable to isolate the effects of every component of the sensitivity index in order to better guide policy actions toward increasing community resilience. Similarly, by using geo-referenced information on health and basic services facilities it is possible to build a more comprehensive measure of sensitivity, which may approximate the supply of health and basic services at the regional level.

References

Abdi, H., L. William and D. Valentin. 2013. "Multiple Factor Analysis: Principal Component Analysis for Multitable and Multiblock Datasets." *Computational Statistics* 5(2): 149-179. doi: 10.1002/wics.1246.

Acemoglu, D., S. Johnson and J. Robinson. 2002. "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution." *Quarterly Journal of Economics* 117(4): 1231-1294. Available at: http://economics.mit.edu/files/4127.

Acemoglu, D., S. Johnson and J. Robinson. 2003. "Disease and Development in Historical Perspective." *Journal of the European Economic Association* 1(2-3): 397-405. Available at: http://scholar.harvard.edu/jrobinson/files/jr_diseasedev.pdf

Barreca, A. 2009. "Climate Change, Humidity, and Mortality in the United States." *Journal of Environmental Economics and Management* 63(1): 19-34.

Birkmann, J. et al. 2013. "Framing Vulnerability, Risk and Societal Responses: The MOVE Framework**."** *Natural Hazards* 67(2): 193-211.

Bosello, F., R. Roson, R.S.J. Tol. 2005. "Economy-Wide Estimates of the Implications of Climate Change: Human Health." Nota di Lavoro 97.2005. Milan, Italy: Fondazione Eni Enrico Mattei.

Burgess, R. et al. 2011. "Weather and Death in India: Mechanisms and Implications of Climate Change." Cambridge, United States: Massachusetts Institute of Technology. Mimeographed document.

Caliendo, M., and S. Kopeining. 2005. "Some Practical Guidance for the Implementation of Propensity Score Matching." IZA Discussion Paper 1558. Bonn, Germany: IZA/Institute for the Study of Labor.

Cameron A.C., and P.K. Trivedi. 1998. *Regression Analysis of Count Data*. Cambridge, United Kingdom: Cambridge University Press.

Cuadras, C. 2012. "Nuevos Metodos de Analisis Multivariante". Capitulos 5 y 6. CMC Editores. Barcelona, Spain.

Curreiro, F.C. et al. 2002. "Temperature and Mortality in 11 Cities of the Eastern United States." *American Journal of Epidemiology* 155(1): 80-87.

Cutter, S.L., B.J. Boruff and W.L. Shirley. 2003. "Social Vulnerability to Environmental Hazards." *Social Science Quarterly* 84(2): 242-261.

Dell, M., B. Jones and B. Olken. 2013. “What Do We Learn from the Weather? The New Climate-Economy Literature.” Cambridge, United States: Harvard University. Forthcoming in *Journal of Economic Literature*. Available at: http://scholar.harvard.edu/files/dell/files/131018_climatejel_manuscript_master_-_post.pdf

Deschênes, O. 2012. “Temperature, Human Health, and Adaptation: A Review of the Empirical Literature.” NBER Working Paper 18345. Cambridge, United States: National Bureau of Economic Research.

Deschênes, O., and M. Greenstone. 2011. “Climate Change, Mortality, and Adaptation: Evidence from Annual Fluctuations in Weather in the US.” *American Economic Journal: Applied Economics* 3(4): 152-85.

Deschênes, O., M. Greenstone and J. Gurryan. 2009. “Climate Change and Birth Weight.” *American Economic Review* 99(2): 211-17.

Deschênes, O., and E. Moretti. 2007. “Extreme Weather Events, Mortality and Migration.” NBER Working Paper 13227. Cambridge, United States: National Bureau of Economic Research.

Derouich, M., A. Boutayeb and E.H. Twizell. 2003. “A Model of Dengue Fever.” *BioMedical Engineering OnLine* 2:4. doi: 10.1186/1475-925X-2-4

Dhiman, R.C., S. Pahwa and P.D. Aditya. 2008. “Climate Change and Health.” Delhi, India: Indian National Institute of Malaria Research (ICMR). Mimeographed document.

Ebi, K.L., R.S. Kovats and B. Menne. 2006. “An Approach for Assessing Human Health Vulnerability and Public Health Interventions to Adapt to Climate Change.” *Environmental Health Perspectives* 114(12): 1930–1934. doi: 10.1289/ehp.8430.

Farmer, P. 1998. “Social Inequalities and Emerging Infectious Diseases.” *Emerging Infectious Diseases* 2(4): 259-269.

Few, R. 2007. “Health and Climatic Hazards: Framing Social Research on Vulnerability, Response and Adaptation.” *Global Environmental Change* 17: 281–295.

Friel, S. et al. 2011. “Urban Health Inequities and the Added Pressure of Climate Change: An Action-Oriented Research Agenda.” *Journal of Urban Health: Bulletin of the New York Academy of Medicine* 88(5): 886-95. doi: 10.1007/s11524-011-9607-0.

Galdo, V., and B. Briceño. 2005. “Evaluating the Impact on Child Mortality of a Water Supply and Sewerage Expansion in Quito: Is Water Enough?” Washington, DC, United States: Inter-American Development Bank, Office of Evaluation and Oversight (OVE).

Galea, S., and D. Vlahov. 2005. “Urban Health: Evidence, Challenges, and Directions.” Annual *Review of Public Health* 26: 341-65.

Galiani, S., P. Gertler and E. Schargrodsky. 2005. “Water for Life: The Impact of the Privatization of Water Services on Child Mortality.” *Journal of Political Economy* 113(1): 83–120.

Githeko, A. et al. 2000. “Climate Change and Vector-Borne Diseases: A Regional Analysis.” *Bulletin of the World Health Organization* 78(9): 1136-1147.

Guerrero Compeán, R. 2013. “Weather and Welfare: Health and Agricultural Impacts of Climate Extremes, Evidence from Mexico.” Working Paper IDB-WP-391. Washington, DC, United States: Inter-American Development Bank.

Guha-Sapir, D., and B. Schimmer. 2005. “Dengue Fever: New Paradigms for a Changing Epidemology.” *Emerging Themes in Epidemiology* 2(1). doi:10.1186/1742-7622-2-1.

Hagenlocher, M. et al. 2013. “Assessing Socioeconomic Vulnerability to Dengue Fever in Cali, Colombia: Statistical vs Expert-Based Modeling.” *International Journal of Health Geographics* 12: 36. doi:10.1186/1476-072X-12-36.

Haines, A., and J.A. Patz. 2004. “Health Effects of Climate Change.” *Journal of the American Medical Association* 291(1): 99–103.

Haines A. et al. 2006. “Climate Change and Human Health: Impacts, Vulnerability, and Mitigation.” *Lancet* 367: 2101–2109.

Heckman, J., H. Ichimura and P.E. Todd. 1998. “Matching as an Econometric Evaluation Estimator.” *Review of Economic Studies* 65: 261-294.

Heinrich, C., A. Maffioli and G. Vázquez. 2010. “A Primer for Applying Propensity- Score Matching.” Technical Note IDB-TN-161. Washington, DC, United States: Inter-American Development Bank.

Hethcote, H. 2000. “The Mathematics of Infectious Diseases.” *SIAM Review* 42(4): 599-653.

Hii, Y.L 2013. “Climate and Dengue Fever: Early Warning Based on Temperature and Rainfall.” Umea, Sweden: Umea University. Doctoral thesis.

Hilbe, J.M. 2011. *Negative Binomial Regression*. Second edition. Cambridge, United Kingdom: Cambridge University Press.

Hirano, K., and G.W. Imbens. 2004. "The Propensity Score with Continuous Treatments." Miami and Berkeley, United States: University of Miami and University of California, Berkeley. Mimeographed document. Available at: http://www.ipc-undp.org/evaluation/aula10-dosagem/Hirano%20%26%20Imbens%20-%20GPS.pdf

Husak, G., J. Michaelsen and C. Funk. 2007. "Use of the Gamma Distribution to Represent Monthly Rainfall in Africa for Drought Monitoring Applications." *International Journal of Climatology* 27: 935-944.

Ikhile, C.I., and P.E. Aifesehi. 2011. "Geographical Distribution of Average Monthly Rainfall in the Western Section of Benin-Owena River Basin, Nigeria." *African Research Review* 5(4): 493-500.

Imai, K., and D.A. Van Dick. 2004. "Causal Inference with General Treatment Regimes: Generalizing the Propensity Score." *Journal of the American Statistical Association* 99: 854-866.

Imbens, G. 2011. "Evaluation Methods with Multivalued Treatments Assuming Selection on Observables." Paper presented at the conference "Mind the Gap," Cuernavaca, Mexico.

Imbens, G., and D. Rubin. 2007. "Matched Sampling for Causal Effect." *International Statistical Review* 75(2): 257-258. doi: 10.1111/j.1751-5823.2007.00015_10.x

Inter-American Development Bank. 2011. *Urban Sustainability in Latin America and the Caribbean.* Washington, United States: Inter-American Development Bank.

Intergovernmental Panel on Climate Change (IPCC). 2007. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. S. Solomon et al., editors. Cambridge, United Kingdom: Cambridge University Press.

Jalan, J., and M. Ravallion. 2003. "Estimating the Benefit Incidence of an Antipoverty Program by Propensity-Score Matching." *Journal of Business & Economic Statistics* 21(1): 19–30.

Johansson, M.A., F. Dominici and G.E. Glass. 2009a. "El Niño Southern Oscillation, Weather, and Dengue Incidence in Puerto Rico, Mexico, and Thailand: A Longitudinal Data Analysis." *PLoS Medicine* 6(11).

http://www.plosmedicine.org/article/info%3Adoi%2F10.1371%2Fjournal.pmed.1000168

----. 2009b. “Local and Global Effects of Climate on Dengue Transmission in Puerto Rico.” *PLoS Neglected Tropical Diseases* 3(2). http://www.plosntds.org/article/info%3Adoi%2F10.1371%2Fjournal.pntd.0000382

Kjellstrom, T. et al. 2007. “Urban Environmental Health Hazards and Health Equity.” *Journal of Urban Health: Bulletin of the New York Academy of Medicine* 84(Supplement 1): 86-97.

Kovats, S. et al. 2000. *Climate Change and Human Health*. Geneva, Switzerland: World Health Organization.

Kovats, S. et al. 2003. “National Assessments of Health Impacts of Climate Change: A Review.” In: A.J. McMichael et al., editors. *Climate Change and Human Health: Risks and Responses*. Geneva, Switzerland: World Health Organization/World Meteorological Organization/United Nations Environment Programme.

Leuven, E., and B. Sianesi. 2003. “PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing.” Available at: http://fmwww.bc.edu/repec/bocode/p/psmatch2.ado

Martcheva, M., and O. Prosper. 2012. “Unstable Dynamics of Vector-Borne Diseases: Modeling through Delay-Differential Equations.” In: V.S.H. Rao and R. Durvasula, editors. *Dynamic Models of Infectious Diseases: Vol 1. Vector Borne Diseases*. New York, United States: Springer.

Medina-Ramón, M., and J. Schwartz. 2007. “Temperature, Temperature Extremes, and Mortality: A Study of Acclimatization and Effect Modification in 50 US Cities.” *Occupational and Environmental Medicine* 64(12): 827–833. doi: 10.1136/oem.2007.033175

Mejia, A. 2013. “Los Desafíos del Agua Potable y El Saneamiento: Informalidad Urbana y Servicios de Agua y Saneamiento.” En: *Equidad e Inclusión Social en América Latina: Acceso Universal al Agua y el Saneamiento*. Serie Reflexiones Sobre Políticas Sociales y Ambientales 2. Bogota, Colombia: Corporacion Andina de Fomento.

Miller, S., S. Yoon and B-K. Yu. 2013. “Vulnerability Indicators of Adaptation to Climate Change and Policy Implications for IDB Projects.” Policy Brief IDB-PB-184. Washington, DC, United States: Inter-American Development Bank.

Molina, O. 2010. “El Impacto de los Cambios Climáticos sobre la Salud en Bolivia: Estimación de Costos y Beneficios hasta el 2100.” Documento de Trabajo 14/2009. La Paz, Bolivia: Instituto de Estudios Avanzados en Desarrollo.

Northridge, M.E. and L. Freeman. 2011. “Urban Planning and Health Equity.” Journal of Urban Health: Bulletin of the New York Academy of Medicine 88(3): 582-597. doi: 10.1007/s11524-011-9558-5

OECD (Organisation for Economic Co-operation and Development). 2008. *Handbook on Constructing Composite Indicators, Methodology and User Guide*. Paris, France: OECD.

PAHO (Panamerican Health Organization). 2013. “Maps and Statistics Portal”. Available at http://www.paho.org/hq/index.php?option=com_content&view=category&layout=blog&id=1221&Itemid=2481”.

Patz, J. et al. 2005. “Impact of Regional Climate Change on Human Health.” *Nature* 438(17): 310-317. doi:10.1038/nature04188

Patz, J.A. et al. 1996. “Global Climate Change and Emerging Infectious Diseases.” *Journal of the American Medical Association* 275(3): 217-223.

Pereda, P., T.A. de Menezes and D. Alves. 2014. Climate Change Impacts on Birth Outcomes in Brazil.” Working Paper IDB-WP-495. Washington, DC, United States: Inter-American Development Bank.

Rao, S. et al. 2007. “The Algebra and Statistics of Generalized Principal Component Analysis.” In: *Visual Communication and Image Processing* 2007. doi: 10.1117/12.707527

Rolandi, C., and P. Schilman. 2012**. “**Linking Global Warming, Metabolic Rate of Hematophagous Vectors, and the Transmission of Infectious Diseases. *Frontiers in Physiology* 3: 75. doi: 10.3389/fphys.2012.00075

Roseghini, W.F.F. et al. 2011. “Dengue Epidemics in Middle-South of Brazil: Climate Constraints (?) and Some Social Aspects.” *Revista Brasileira de Climatologia* 9(July-December): 94-101. http://ojs.c3sl.ufpr.br/ojs2/index.php/revistaabclima/article/viewFile/27522/18336

Rosenbaum, P., and D. Rubin. 1983. “The Central Role of the Propensity Score in Observational Studies for Causal Effects.” Biometrika 70, 41–55.

Ruprah, I.J. 2009. “The Housing Gap in Latin America: 1995–2015.” Washington, DC, United States: Inter-American Development Bank. Manuscript.

Suhaila, J. et al. 2011. "Introducing the Mixed Distribution in Fitting Rainfall Data." *Open Journal of Modern Hydrology* 11(1): 11-22.

Sutherst, R. 2004. "Global Change and Human Vulnerability to Vector-Borne Diseases." *Clinical Microbiology Reviews* 17(1): 136-173. doi: 10.1128/CMR.17.1.136-173.2004.

Tarazona, M., and J. Gallegos. 2011. "Recent Trends in Disaster Impacts on Child Welfare and Development 1999-2009." Global Assessment Report on Disaster Risk Reduction. London, United Kingdom: Children in a Changing Climate. Available at: http://www.preventionweb.net/preventionweb-files/english/hyogo/gar/2011/en/bgdocs/Tarazona_&_Gallegos_2010.pdf

Tol, S.J.R. 2002a. "Estimates of Damage Costs of Climate Change. Part I: Benchmark Estimates." *Environmental and Resource Economics* 21(1): 47-73. doi: 10.1023/A:1014500930521.

----. 2002b. "Estimates of Damage Costs of Climate Change. Part II: Dynamic Estimates." *Environmental and Resource Economics* 21: 135–160. doi:10.1023/A:1014539414591.

Thurston, G., K. Ito. 2001. "Epidemiological Studies of Acute Ozone Exposures and Mortality." *Journal of Environmental Epidemiology* 11: 286-294

Valencia, M. 2014. "The Health Impacts of Severe Climate Shocks in Colombia." Working Paper IDB-WP-498. Washington, DC, United States: Inter-American Development Bank.

Vincent, K., and J.M. Sutherland. 2013. "A Review of Methods for Deriving an Index of Socioeconomic Status in British Columbia." Vancouver, Canada: University of British Columbia, Centre for Health Services and Policy Research. Available at: http://healthcarefunding.ca/files/2013/04/Review-of-Methods-for-SES-Index-for-BC.pdf

Vlahov, D. et al. 2007. "Urban as a Determinant of Health." *Journal of Urban Health: Bulletin of the New York Academy of Medicine* 84(Supplement 1): 16-26. doi: 10.1007/s11524-007-9169-3.

Wooldridge, J. 2010. *Econometric Analysis of Cross Section and Panel Data*. Second edition. Cambridge, United States: MIT Press.

Wooldridge, J., and A. Semykina. 2010. "Estimating Panel Data Models in the Presence of Endogeneity and Selection." *Journal of Econometrics* 157: 375-380.

World Health Organization (WHO). 2003. "Methods of Assessing Human Health Vulnerability and Public Health Adaptation to Climate Change." Health and Global Environmental

Change, Series 1. Copenhagen, Denmark: World Health Organization Europe. Available at: http://www.euro.who.int/__data/assets/pdf_file/0009/91098/E81923.pdf

----. 2010. *Hidden Cities: Unmasking and Overcoming Health Inequities in Urban Settings*. Geneva, Switzerland: World Health Organization Centre for Health Development and the United Nations Settlements Program. http://whqlibdoc.who.int/publications/2010/9789241548038_eng.pdf

Table 1. Dengue and Malaria: Average Number of Cases and Incidence by Country, 1995-2010

Country	Criteria	Malaria		Dengue	
		Mean	**Std. Dev.**	**Mean**	**Std. Dev.**
Belize	Number of registered cases of [Cases]	2,247	2,637	315	660
	Annual Parasitic Incidence (per 100 thousand pop surv)	0.012	0.012	0.068	0.110
Bolivia	Number of registered cases of [Cases]	29,294	20,753	8,102	20,641
	Annual Parasitic Incidence (per 100 thousand pop surv)	0.013	0.006	0.230	0.245
Brazil	Number of registered cases of [Cases]	444,814	104,200	404,555	254,967
	Annual Parasitic Incidence (per 100 thousand pop surv)	0.032	0.028	0.258	0.142
Colombia	Number of registered cases of [Cases]	139,672	43,514	49,388	33,816
	Annual Parasitic Incidence (per 100 thousand pop surv)	0.021	0.015	0.228	0.157
Costa Rica	Number of registered cases of [Cases]	2,338	1,781	13,092	10,570
	Annual Parasitic Incidence (per 100 thousand pop surv)	0.004	0.003	0.560	0.332
Dominican Rep.	Number of registered cases of [Cases]	1,961	1,023	4,232	3,470
	Annual Parasitic Incidence (per 100 thousand pop surv)	0.003	0.002	0.064	0.035
El Salvador	Number of registered cases of [Cases]	1,006	1,748	9,371	7,825
	Annual Parasitic Incidence (per 100 thousand pop surv)	0.0005	0.0004	0.188	0.107
Honduras	Number of registered cases of [Cases]	28,295	22,303	21,128	14,580
	Annual Parasitic Incidence (per 100 thousand pop surv)	0.009	0.004	0.338	0.247
Mexico	Number of registered cases of [Cases]	4,712	3,323	45,246	61,944
	Annual Parasitic Incidence (per 100 thousand pop surv)	0.0008	0.0011	0.034	0.042
Nicaragua	Number of registered cases of [Cases]	19,498	24,070	5,859	6,056
	Annual Parasitic Incidence (per 100 thousand pop surv)	0.009	0.005	0.072	0.075
Paraguay	Number of registered cases of [Cases]	2,108	2,789	6,347	10,054
	Annual Parasitic Incidence (per 100 thousand pop surv)	0.0013	0.0014	131.430	182.517
Peru	Number of registered cases of [Cases]	115,820	72,452	7,683	6,300
	Annual Parasitic Incidence (per 100 thousand pop surv)	0.016	0.008	0.034	0.022
Venezuela	Number of registered cases of [Cases]	30,965	11,591	46,368	28,880
	Annual Parasitic Incidence (per 100 thousand pop surv)	0.024	0.019	0.200	0.099

Source: PAHO.

Table 2. Climatic Limiting Conditions, by Vector-Borne Disease

Disease	Pathogen	Vector	Minimum Temperature ºC	Maximum Temeprature ºC	Minimum temperature for vector ºC
Malaria	Plasmodium falciparum Plasmodium vivax	Anopheles	8 – 10	33 - 39	8 - 1 (biological activity)
Dengue	Dengue virus	Aedes	14- 15	33 - 39	8 - 10 (biological activity)
Chagas	Trupanosomacruzi	Triatomina (kissing bugs)	18	38	2 -6 (survival), 20 (biological activity)
Lyme Disease	Borreliaburdoferi	Ixodes ticks	Not determined	Not determined	5 – 8

Source: Dhiman, Pahwa and Aditya (2010).

Table 3. Variable Description and Sources

Variable Type	Description	Sources	Administrative / Geographical
Dengue and malaria outbreaks	Occurrence of epidemic outbreak, by administrative layer Intensity of an outbreak, by administrative layer	Health Maps	3rd Administrative
Socioeconomic Variables: Costa Rica	Selected indicators reflecting local assets, by administrative level	INEC, RADETAM-CELADE	3rd, 4th Administrative
Socioeconomic Variables: Mexico	Selected indicators reflecting local assets, by administrative level	INEGI	3rd, 4th Administrative
Socioeconomic Variables: Brazil	Selected indicators reflecting local assets, by administrative level	IPUMS	3rd Administrative and Household level information
Socioeconomic Variables: Nicaragua	Selected indicators reflecting local assets, by administrative level	IPUMS	3rd Administrative and Household level information
Socioeconomic Variables: Colombia	Selected indicators reflecting local assets, by administrative level	IPUMS	3rd, 4th Administrative
Rainfall	Monthly, seasonal or annual rainfall, by geographical region Statistical moments for rainfall, by geographical region	NASA's Tropical Rainfall Measuring Mission (TRIMM) Satellite Databases, GEOPORTAL	0.25ºx0.25º Pixel Resolution
Temperature	Monthly temperature average , by geographical region	NOOA's Satellite Databases. Retrieved from IRI – University of Columbia climate repository GEOPORTAL	0.5ºx0.5º Pixel Resolution

Source: Authors' compilation.

Table 4. Number of Reported Cases by Year and Country

	Costa Rica	**Mexico**	**Dominican Republic**	**Brazil**	**Colombia**	**Nicaragua**
2005	0	0	2	0	0	0
2006	1	1	1	1	0	0
2007	0	1	1	35	0	0
2008	0	13	1	24	0	0
2009	0	30	1	43	1	3
2010	0	9	0	37	10	1
2011	0	3	0	14	1	2
2012	2	29	6	24	19	5
2013	2	58	5	26	22	8
Total	**5**	**144**	**17**	**204**	**53**	**19**
Number of 3rd admin. layer locations with more than one reported outbreak	2	20	3	108	29	8
Total number of 3rd admin. layer locations with reported outbreaks	3	140	15	204	65	19

Source: Health Maps.

Table 5. Intensity Statistics of Epidemic Outbreaks, by Country

	Injured - Dengue		Injured - Malaria	
	Mean	Std. Dev	Mean	Std. Dev
Costa Rica	418	342	77.5	91.21
Mexico	149	347	5.67	3.05
Brazil	2790	12300	17.67	13.31
Colombia	411	868	46	57
Nicaragua	214	305	0	0
Dominican Republic	104	259	20.51	0.7

Source: Health Maps.

Table 6. Monthly Distribution of Reported Cases, by Country

	Dengue						Malaria					
	Mexico	Costa Rica	Dom. Republic	Colombia	Nicaragua	Brazil	Mexico	Costa Rica	Dom. Republic	Colombia	Nicaragua	Brazil
January	11		1	2		8				2		
February	15		1	1		36			1	1		
March	3	1	4	4	1	24		1		1		1
April	4		1	8		21			1			
May	12		1	12	1	12			1	2		
June	3		1	2	1	9						
July	39		2	5	5	25				3		
August	15	2	1	2	1	5						1
September	25		1	3	4	22						
October	4		1	9	3	14				1		
November	5		1	1		11	3	1		2		
December	4			4	3	15						1

Source: Health Maps.

Table 7. Number of Available Climatic Information Pixels, by Country

Country	**Number of pixels**	**Number of pixels actually available**
Mexico	2,764	1,071
Nicaragua	1,73	
El Salvador	28	
Honduras	150	
Haiti	35	
Dominican Republic	63	48
Costa Rica	69	39
Bolivia	1,468	
Brazil	10,711	2,754
Paraguay	562	
Peru	1,706	
Colombia	1,480	
Total	**19,209**	**3,912**

Source: Authors' compilation.

Table 8. Sources of Socioeconomic Information, by Country

Country	Type of Survey	Name	Year	Admin. Level Reached	Source	Primary Source
Mexico	Census	Censo de Población y Vivienda	2010	3rd and 4th Administrative Level	INEGI	INEGI
Mexico	Census	Censo Nacional de Población y Vivienda	2000	3rd Administrative Level	INEGI	INEGI
Costa Rica	Census	Censo Nacional de Población y Vivienda	2011	3rd and 4th Administrative Level	RADETAM	INEC
Costa Rica	Census	Censo Nacional de Población y Vivienda	2000	3rd and 4th Administrative Level	RADETAM	INEC
Dominican Republic	Census	Censo	2010	3rd and 4th Administrative Level	RADETAM	ONE
Colombia	Census	Censo	2007	3rd and 4th Administrative Level	IPUMS	DANE
Nicaragua	Census	Censo	2005	3rd and 4th Administrative Level	IPUMS	INED
Brazil	Census Sample		2010	3rd Administrative and Household Level	IPUMS	IBGE
Brazil	Census Sample	Censo	2000	3rd Administrative and Household Level	IPUMS	IBGE

Source: INEGI, INEC, ONE, RADETAM, World Bank Microdata Catalogue and IPUMS.

Table 9. Pre-Treatment Differences, by Country

	Brazil							Colombia						
	Treatment			Control			Normalized Differences	Treatment			Control			Normalized Differences
Variable	N	Mean	Std. Deviation	N	Mean	Std. Deviation		N	Mean	Std. Deviation	N	Mean	Std. Deviation	
sewage	82	0.744	0.251	1414	0.509	0.31	0.59	34	0.850	0.120	206	0.682	0.215	0.682
unemployment	82	0.051	0.02	1414	0.069	0.034	-0.474	34	0.073	0.023	206	0.047	0.032	0.663
electricity	82	0.987	0.011	1414	0.969	0.052	0.338	34	0.935	0.083	206	0.859	0.135	0.483
literacy	82	0.089	0.074	1414	0.14	0.087	-0.448	34	0.145	0.074	206	0.206	0.069	-0.593
share_pop<14years	82	0.226	0.031	1414	0.258	0.049	-0.549	34	0.302	0.054	206	0.337	0.047	-0.483
total_population	82	15414.04	30723.21	1414	6560.05	21399.8	0.236	34	32263.2	26519.07	206	7605.7	10665.36	0.863
latitude	203	-48.366	6.077	2659	-48.61	5.795	0.029	34	-74.721	1.486	206	-74.972	1.452	0.121
longitude	203	-18.262	7.158	2659	-16.43	8.238	-0.168	34	7.574	2.697	206	6.215	2.452	0.373
surface	203	352520.2	770314.8	2659	282331	847450.6	0.061	34	1392.7	1537.5	206	1258.4	2027.9	0.053
temperature	203	297.089	3.254	2659	295.743	66.593	0.02	34	298.309	3.580	206	297.038	4.660	0.216
precipitation	203	108.544	104.256	2659	113.891	110.373	-0.035	34	180.378	43.798	206	217.076	79.201	-0.405
density	82	0.168	0.275	389	0.059	0.151	0.348	34	51.296	52.771	206	16.737	28.241	0.577

Table 9., continued

	Mexico							Nicaragua						
	Treatment			Control			Normalized Differences	Treatment			Control			Normalized Differences
Variable	N	Mean	Std. Deviation	N	Mean	Std. Deviation		N	Mean	Std. Deviation	N	Mean	Std. Deviation	
sewage	141	0.886	0.148	2368	0.739	0.246	0.512	19	0.742	0.172	37	0.092	0.098	3.291
unemployment	141	0.041	0.015	2368	0.041	0.033	-0.021	19	0.037	0.009	37	0.019	0.013	1.089
electricity	141	0.972	0.024	2368	0.949	0.064	0.332	19	0.836	0.150	37	0.483	0.272	1.137
literacy	141	0.066	0.042	2368	0.123	0.074	-0.668	19	0.169	0.062	37	0.308	0.095	-1.221
share_pop<14years	141	0.041	0.015	2368	0.041	0.033	-0.021	19	0.345	0.034	37	0.415	0.055	-1.084
total_population	141	291640.8	353121.5	2368	39911.9	120520.1	0.675	19	26720.1	35947.92	37	4530	2350.875	0.616
latitude	99	-97.294	7.038	1015	-100.346	5.236	0.348	19	-85.783	0.796	37	-85.368	1.073	-0.310
longitude	99	21.552	3.139	1015	21.592	4.128	-0.008	19	12.664	0.529	37	12.841	0.760	-0.191
surface	99	3520000	4760000	1015	1570000	2940000	0.349	19	1011.5	988.3	37	2024.9	2079.4	-0.440
temperature	99	276.565	205.988	1015	287.251	152.173	-0.042	19	299.895	1.096	37	159.517	846.241	0.166
precipitation	99	54.126	86.612	1015	51.871	86.883	0.018	19	163.783	22.787	37	163.686	36.890	0.002
density	99	0.001	0.002	1015	0	0.001	0.267	19	36.490	50.563	37	5.706	8.713	0.600

Table 9., continued

	Costa Rica							Dominican Republic						
	Treatment			**Control**			**Normalized Differences**	**Treatment**			**Control**			**Normalized Differences**
Variable	**N**	**Mean**	**Std. Deviation**	**N**	**Mean**	**Std. Deviation**		**N**	**Mean**	**Std. Deviation**	**N**	**Mean**	**Std. Deviation**	
sewage	4	0.931	0.037	78	0.948	0.056	-0.251	18	0.428	0.128	148	0.626	0.174	-0.92
unemployment	4	0.043	0.009	78	0.032	0.008	0.994	18	0.065	0.018	148	0.088	0.03	-0.665
electricity	4	0.973	0.023	78	0.985	0.026	-0.323	18	0.07	0.059	148	0.125	0.14	-0.366
literacy	4	0.144	0.094	78	0.061	0.079	0.672	18	0.206	0.025	148	0.246	0.07	-0.534
share_pop<14years	4	0.273	0.043	78	0.251	0.036	0.384	18	0.285	0.017	148	0.313	0.035	-0.699
total_population	4	123249.8	115904	78	50040.1	45311.81	0.588	18	60359	40524.15	148	22659	55392.35	0.549
latitude	3	-83.548	0.676	37	-84.328	0.82	0.734	7	-69.655	1.274	44	-70.75	0.663	0.763
longitude	3	9.703	0.12	37	9.979	0.637	-0.425	7	18.978	0.507	44	18.923	0.498	0.077
surface	3	1.01E+09	7.58E+08	37	1.4E+09	8.69E+08	-0.321	7	542.5	162.7	44	274.1	208	1.016
temperature	3	297.488	1.105	37	298.502	2.759	-0.341	7	299.1	1.665	44	298.29	2.385	0.278
precipitation	3	142.362	174.088	37	130.631	168.334	0.048	7	67.683	75.548	44	65.18	82.383	0.022
density	3	0	0	37	0	0	-0.062	7	92.021	55.972	44	103.5	142.198	-0.075

Source: Authors' calculations.

Table 10. National Level Covariates Included

	GDP per capita 2012	Health Expenditure per capital 2012	Share of Health Expenditure in GDP	Average of Private Share in Health Expenditure	Private Out of Pocket Share on Health Expenditure
	USD in PPP	**USD in PPP**	**%**	**%**	**%**
Brazil	11339.52	1042.73	8.90	56.23	58.17
Costa Rica	9396.45	1329.81	10.87	31.59	89.28
Dominican Republic	5736.44	529.07	5.36	51.33	78.83
Mexico	9747.46	940.10	6.16	52.63	92.71

Source: WHO.

Table 12. Summary Statistics and Correlations Matrices Simple Propensity Score (Logit)

Variable	Obs	Mean	Std. Dev.	Min	Max
Gross_knots	1,914	0.1031	0.1833	0.0000	0.9917
Gross_bins	1,914	0.1042	0.1910	0.0000	0.9963
Gross-thresholds	1,914	0.1035	0.1643	0.0000	0.9670
PTE_SOC_Knots	1,914	0.0869	0.1667	0.0000	0.9759
PTE_CLIM_Knots	1,914	0.1126	0.1984	0.0000	0.8868
PTE_SOC_Bins	1,914	0.0760	0.1586	0.0000	0.9489
PTE_CLIM_Bins	1,914	0.1160	0.2062	0.0000	0.9150
PTE_SOC_Thresholds	1,914	0.1562	0.2297	0.0000	0.9943
PTE_CLIM_Thresholds	1,914	0.2058	0.2989	0.0006	0.9641

Source: Authors’ calculations.

Table 11. Propensity Score Model for Epidemic Dengue and Malaria Outbreaks

Model	**PHYSICAL THRESHOLDS AND LEVELS**		**CLIMATIC BINS**		**DISTRIBUTION THRESHOLDS**
	MODEL 1		**MODEL 2**		**MODEL 3**
N	**1831**		**1831**		**1831**
AIC	**744.218**		**749.236**		**794.240**
BIC	**898.571**		**991.791**		**887.954**
Pseudo Likelihood	**-344.109**		**-330.618**		**-380.120**
Chi2	**0.000**		**0.000**		**0.000**
LROC	**0.894**		**0.904**		**0.855**
Pseudo - R2	**0.349**		**0.375**		**0.281**
threshold > 20º C	**5.782**	b_2_precipitation	**0.079**	dummy threshold perc. 95 temperature	**1.903**
threshold > 100 mm	**0.466**	b_3_precipitation	0.029	dummy threshold perc. 95 precipitation	0.278
threshold > 25º C	**-1015.985**	b_6_precipitation	**0.163**	dummy threshold perc. 90 temperature	0.001
threshold > 150 mm	**-80.774**	b_11_precipitation	**0.085**	dummy threshold perc. 90 precipitation	**0.537**
temperature continous > 20ºC	**-3.488**	b_17_precipitation	**-0.059**	dummy threshold perc. 75 temperature	0.201
temperature continous > 25ºC	**0.000**	b_18_precipitation	0.032	dummy threshold perc. 75 precipitation	-1.096
precipitation continous >100 mm	**-0.520**				
precipitation continous >150 mm	**-0.001**				
temperature continous > 20ºC squared	**3.442**				
temperature continous > 25ºC squared	0.000				
pp continous >100 mm squared	**0.521**				
pp continous >150 mm squared	**0.001**				
year mean temperature	**0.021**	year mean temperature	**0.033**		
year mean precipitation	0.011	year mean precipitation	0.008		
summer mean precipitation	0.005	summer mean precipitation	-0.002		
summer mean temperature	**-0.022**	summer mean temperature	**-0.347**		
population logs	**1.115**	population logs	**1.187**	population logs	**1.116**
density logs	0.207	density logs	0.278	density logs	**-0.371**
share_pop<14years logs	**0.261**	share_pop<14years logs	0.099	share_pop<14years logs	**1.177**
illiteracy logs	-0.529	illiteracy logs	**-0.734**	illiteracy logs	-0.604
electricity logs	0.804	electricity logs	0.633	electricity logs	**1.100**
unemployment logs	0.487	unemployment logs	**0.581**	unemployment logs	**1.152**
water supply logs	**1.955**	water supply logs	**1.722**	water supply logs	**1.666**
GDP per capita logs	3.195	GDP per capita logs	3.853	GDP per capita logs	8.428
health expenditure per capita logs	**-6.345**	health expenditure per capita logs	**-7.429**	health expenditure per capita logs	**-12.018**
latitude	0.033	latitude	-0.016	latitude	-0.006
longitude	**0.111**	longitude	**0.073**	longitude	**0.106**
constant	4.815	constant	12.373	constant	1.246

Source: Authors' calculations.

Table 12. Summary of the Balancing Conditions, by Propensity Score Model

MODEL 1. PHYSICAL THRESHOLDS AND LEVELS

Climatic Regressors		Socioeconomic Regressors	
Balanced	**Unbalanced**	**Balanced**	**Unbalanced**
threshold > 20º C		population logs	
threshold > 100 mm		share_pop<14years logs	density logs
threshold > 25º C		illiteracy logs	
threshold > 150 mm		electricity logs	
temperature continuous > 20ºC		unemployment logs	
precipitation continuous >100 mm	temperature continuous > 25ºC	water supply logs	
precipitation continuous >150 mm		GDP per capita logs	
temperature continuous > 20ºC squared		health expenditure per capita logs	
temperature continuous > 25ºC squared			
pp continuous >100 mm squared			
pp continuous >150 mm squared			
year mean temperature			
year mean precipitation			
summer mean precipitation			
summer mean temperature			
latitude			
longitude			

MODEL 2. CLIMATIC BINS

Climatic Regressors		Socioeconomic Regressors	
Balanced	**Unbalanced**	**Balanced**	**Unbalanced**
b_2_precipitation		population logs	
b_3_precipitation		density logs	
b_6_precipitation		share_pop<14years logs	
b_11_precipitation		electricity logs	illiteracy logs
b_17_precipitation		unemployment logs	
b_18_precipitation		water supply logs	
year mean temperature		GDP per capita logs	
year mean precipitation		health expenditure per capita logs	
summer mean precipitation			
summer mean temperature			
latitude			
longitude			
constant			

MODEL 3. DISTRIBUTION THRESHOLDS

Climatic Regressors		Socioeconomic Regressors	
Balanced	**Unbalanced**	**Balanced**	**Unbalanced**
dummy threshold perc. 95 temperature	latitude	population logs	
dummy threshold perc. 95 precipitation	longitude	density logs	
dummy threshold perc. 90 temperature		share_pop<14years logs	
dummy threshold perc. 90 precipitation		illiteracy logs	
dummy threshold perc. 75 temperature		electricity logs	
dummy threshold perc. 75 precipitation		unemployment logs	
		water supply logs	
		GDP per capita logs	
			health expenditure per capita logs

Table 13. Marginal Effects of Selected Variables

	SIMPLE PROPENSITY SCORE (LOGIT)								
	PHYSICAL THRESHOLDS AND LEVELS			CLIMATIC BINS			DISTRIBUTION THRESHOLDS		
	MODEL 1			MODEL 2			MODEL 3		
	Odds ratios	Average Marginal Effects		Odds ratios	Average Marginal Effects		Odds ratios	Average Marginal Effects	
		Marginal Effects	Semi-elasticities		Marginal Effects	Semi-elasticities		Marginal Effects	Semi-elasticities
ln_density	**0.793**		0.012	0.776		0.009	**0.690**		**0.022**
ln_por_pop<14	0.929		0.008	0.823		0.006	**0.314**		**-0.069**
ln_water_supply	**6.523**		**0.105**	**5.205**		**0.063**	**5.723**		**0.098**
ln_health_exp	**0.058**		**-0.344**	**-0.041**		**-0.418**	**2.845**		-0.706
ln_pop_total	**3.278**		**0.060**	**3.436**		**0.059**			**0.066**
threshold > 20º C	**40.494**	**0.313**	**40.247**						
threshold > 100 mm	**1.547**	**0.025**	**2.025**						
threshold > 25º C	**0.000**	**-55.059**	**-24.642**						
threshold > 150 mm	**0.000**	**-4.377**	**-1.391**						
temperature continuous > 20ºC	**-0.027**	**-0.189**	**-40.324**						
temperature continuous > 25ºC	**1.000**	**0.000**	**-0.170**						
precipitation continuous >100 mm	**-0.621**	**-0.028**	**-2.716**						
precipitation continuous >150 mm	**0.999**	**0.000**	**-0.473**						
temperature continuous > 20ºC squared	**-36.650**	**0.187**	**25.079**						
temperature continuous > 25ºC squared	**1.000**	0.000	0.045						
pp continuous >100 mm squared	**1.604**	**0.000**	**2.277**						
pp continuous >150 mm squared	**1.001**	**0.000**	**0.324**						
dummy threshold perc. 95 temperature							**6.224**	**0.112**	**0.687**
dummy threshold perc. 95 precipitation							1.308	0.016	0.099
dummy threshold perc. 90 temperature							1.079	0.006	0.065
dummy threshold perc. 90 precipitation							**1.745**	**0.032**	**0.367**
dummy threshold perc. 75 temperature							1.249	0.012	0.340
dummy threshold perc. 75 precipitation							0.524	**0.064**	**-1.849**

Source: Authors' calculations.

Table 14. Generalized Propensity Score (Negative Binomial) Detailed Results

Model	**PHYSICAL THRESHOLDS AND LEVELS**		**CLIMATIC BINS**		**DISTRIBUTION THRESHOLDS**
	MODEL 1		**MODEL 2**		**MODEL 3**
N	**1,826**		**1,826**		**1,831**
AIC	**2443.406**		**2424.965**		**794.240**
BIC	**2597.683**		**2667.399**		**887.954**
Pseudo Likelihood	**-1193.703**		**-330.618**		**-380.120**
Clusters	**104**		**104**		**104**
Chi2	**0.000**		**0.000**		**0.000**
Pseudo - R2	**0.069**		**0.088**		**0.040**
threshold > 20º C	4.972	b_3_temperature	**0.499**	dummy threshold perc. 95 temperature	**0.693**
threshold > 100 mm	**1.043**	b_4_temperature	**0.489**	dummy threshold perc. 95 precipitation	0.643
threshold > 25º C	-813.280	b_5_temperature	**0.505**	dummy threshold perc. 90 temperature	**-1.671**
threshold > 150 mm	**-191.548**	b_6_temperature	**0.507**	dummy threshold perc. 90 precipitation	**1.987**
temperature continuous > 20ºC	-2.966	b_7_temperature	**0.491**	dummy threshold perc. 75 temperature	-0.127
temperature continuous > 25ºC	**0.000**	b_8_temperature	**0.522**	dummy threshold perc. 75 precipitation	**2.319**
precipitation continuous >100 mm	**-1.521**	b_9_temperature	**0.485**		
precipitation continuous >150 mm	**0.001**	b_10_temperature	**0.547**		
temperature continuous > 20ºC squared	2.749	b_2_precipitation	**0.183**		
temperature continuous > 25ºC squared	0.000	b_6_precipitation	**0.193**		
pp continuous >100 mm squared	**1.482**	b_7_precipitation	**0.450**		
pp continuous >150 mm squared	**-0.001**	b_11_precipitation	**0.322**		
		b_14_precipitation	**-0.328**		
		b_16_precipitation	**0.581**		
year mean temperature	0.048	year mean temperature	**0.120**		
year mean precipitation	0.027	year mean precipitation	0.008		
summer mean precipitation	0.036	summer mean precipitation	-0.002		
summer mean temperature	-0.049	summer mean temperature	**-0.133**		
density logs	0.248	density logs	0.278	density logs	0.209
share_pop<14years logs	-0.845	share_pop<14years logs	**-1.674**	share_pop<14years logs	0.378
illiteracy logs	**-3.372**	illiteracy logs	**-2.741**	illiteracy logs	**-2.799**
electricity logs	4.112	electricity logs	0.633		
unemployment logs	**2.024**	unemployment logs	**2.522**	unemployment logs	**2.177**
water supply logs	**1.640**	water supply logs	**4.219**	water supply logs	**4.228**
GDP per capita logs	4.989	GDP per capita logs	3.853	GDP per capita logs	**-54.047**
health expenditure per capita logs	**-12.493**	health expenditure per capita logs	**-11.049**	health expenditure per capita logs	**-35.662**
latitude	**0.141**	latitude	0.055	latitude	0.035
longitude	**0.215**	longitude	-0.045	longitude	**0.218**
constant	32.295	constant	12.373	constant	**206.918**
alpha	**47.294**	alpha	38.835	alpha	67.4597
ln_alpha	**3.856**	ln_alpha	**3.659**	ln_alpha	**4.212**
población_total	exposure	población_total	exposure	población_total	exposure

Source: Authors' calculations.

Table 15. Sensitivity Index: Components, Brazil 2000

Indicator	PC1 – SUS	
Bathrooms	0.333	Less Susceptible
Average Years of School	0.328	
Refrigerator	0.294	
Phone	0.292	
Automobile	0.261	
High School Attainment	0.260	
Water Supply	0.259	
TV	0.258	
Rooms In House	0.249	
Electricity	0.236	
Literacy	0.211	
Radio	0.152	
Unpaid Worker	0.130	
Sewage System	0.115	
Employed	0.109	
Bedrooms	0.097	
HH-Married With children	0.013	
HH-Married No Children	0.003	
Wage/Salary Worker	-0.001	More Susceptible
Female	-0.005	
HH-Ext. Fam., Relatives Only	-0.006	
Age	-0.010	
HH-Family and Non-rel	-0.022	
HH-One Person	-0.030	
Child Survival Rate	-0.053	
Disabilities	-0.060	
Number of Children	-0.094	
Self-employed	-0.109	
No Trash Service	-0.148	

Indicator	PC2 – LoR	
Employed	0.533	More Resilient
Unpaid Worker	0.370	
Wage/Salary Worker	0.244	
Child Survival Rate	0.106	
High School Attainment	0.092	
HH-Married With children	0.079	
Average Years of School	0.071	
No Trash Service	0.056	
HH-One Person	0.028	
HH-Married No Children	0.024	
Literacy	0.008	
Age	-0.010	Less Resilient
Sewage System	-0.019	
HH-Ext. Fam., Relatives Only	-0.025	
Automobile	-0.028	
Radio	-0.043	
Phone	-0.071	
Bedrooms	-0.084	
Disability	-0.085	
HH-Family and Non-rel	-0.094	
Rooms In House	-0.099	
Water Supply	-0.101	
Refrigerator	-0.105	
TV	-0.105	
Bathrooms	-0.106	
Electricity	-0.111	
Number of Children	-0.166	
Female	-0.270	
Self-employed	-0.533	

Source: Authors' calculations.

Table 16. Average Treatment on the Treated, by Propensity Score

			ORIGINAL PROPENSITY			
N = 1,720	**Variable**	**Sample**	**Treated**	**Controls**	**S.E.**	**T-stat**
MODEL 1. CLIMATIC BINS	avg_si	Unmatched	-2.349	-1.059	1.000	
		ATT	-2.349	-3.202	1.148	0.740
	sd_si	Unmatched	11.272	10.785	0.390	
		ATT	11.272	10.502	0.653	1.180
	p50_si	Unmatched	-2.135	-0.970	1.079	
		ATT	-2.135	-3.056	1.211	0.760
	si_ratio75_25	Unmatched	0.154	-0.670	1.322	
		ATT	0.154	-0.625	3.147	0.250
	si_ratio90_10	Unmatched	-0.493	10.719	35.817	
		ATT	-0.493	11.333	86.206	-0.140
N = 1,720	**Variable**	**Sample**	**Treated**	**Controls**	**S.E.**	**T-stat**
MODELS 2. CLIMATIC KNOTS	avg_si	Unmatched	-2.349	-1.059	1.000	
		ATT	-2.349	-2.473	1.172	0.110
	sd_si	Unmatched	11.272	10.785	0.390	
		ATT	11.272	10.656	0.641	0.960
	p50_si	Unmatched	-2.135	-0.970	1.079	
		ATT	-2.135	-2.289	1.215	0.130
	si_ratio75_25	Unmatched	0.154	-0.670	1.322	
		ATT	0.154	-0.638	0.792	0.41
	si_ratio90_10	Unmatched	-0.493	10.719	35.817	
		ATT	-0.493	11.888	-12.382	-0.16
N = 1,720	**Variable**	**Sample**	**Treated**	**Controls**	**S.E.**	**T-stat**
MODEL 3. CLIMATIC THRESHOLDS	avg_si	Unmatched	-2.349	-1.059	1.000	
		ATT	-2.337	-3.238	1.142	0.790
	sd_si	Unmatched	11.272	0.487	0.390	
		ATT	11.286	10.968	0.610	0.520
	p50_si	Unmatched	-2.135	-1.165	1.079	
		ATT	2.123	-3.143	1.019	0.850
	si_ratio75_25	Unmatched	0.154	-0.670	1.322	
		ATT	0.154	-0.826	1.788	0.550
	si_ratio90_10	Unmatched	-0.493	10.719	35.817	
		ATT	-0.493	81.188	65.659	-1.240

Source: Authors' calculations.

Table 17a. Test for Common Trends Hypothesis

	Obs.	**Mean**	**SD**
Treatment	129	1.024	0.659
Control	1,340	2.473	0.361
combined	1,469	2.346	0.334
diff		-1.449	1.182
Ho: Treatment and Control Differences are Equal			
t	**0.2204**		
Degrees of freedom	1467		

Table 17b. ATTT of Differences-in-Differences

Sensitivity Index Moments	**Treated**	**Control**	**Differences**	**t-statistic**
Model 1. Knots				
Standard Deviation	8.35928849	8.82236041	-0.463071927	-1.19
Mean	1.62066972	-1.40183085	-0.218838868	-0.27
Model 2. Climatic Bins				
Standard Deviation	8.35928849	8.79665757	-0.437369084	-0.95
Mean	-1.62066972	-1.54943389	-0.071235837	-0.07
Model 3. Distribution Thresholds				
Standard Deviation	8.35928849	8.82788355	-0.468595062	-1.39
Mean	-1.62066972	-1.18809799	-0.432571733	-0.58

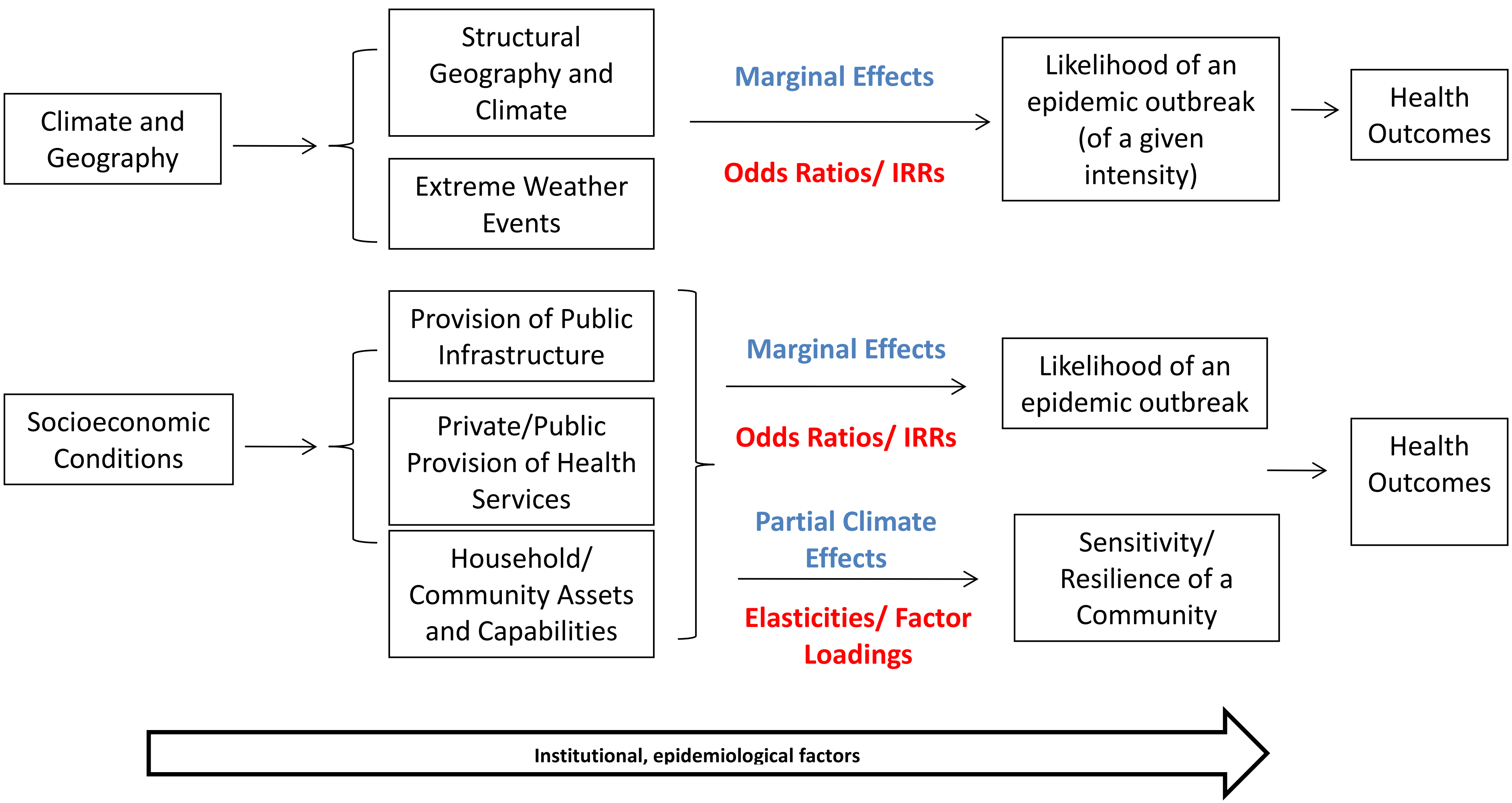
Source: Authors' calculations.

Table 18. Most and Least Vulnerable Cities with More than 50 Thousand Inhabitants

BOTTOM 30 MOST VULNERABLE CITIES			TOP 30 MOST RESILLIENT CITIES		
COUNTRY	STATE	MUNICPALITY	COUNTRY	STATE	MUNICPALITY
Brazil	SÃO PAULO	São Paulo	Costa Rica	Heredia	Sarapiquí
Brazil	Rio de Janeiro	Rio de Janeiro	Costa Rica	Limon	Limon
Brazil	Distrito Federal	Brasilia	Costa Rica	Limón	Pococí
Mexico	Tamaulipas	Reynosa	Costa Rica	Cartago	Turrialba
Mexico	Quintana Roo	Benito Juarez	Costa Rica	Limón	Siquirres
Mexico	Tabasco	Centro-VillaHermosa	Costa Rica	Guanacaste	Nicoya
Mexico	Sonora	Guaymas	Costa Rica	Puntarenas	Puntarenas
Mexico	Michoacan	Morelia	Costa Rica	Alajuela	San Carlos
Costa Rica	Alajuela	San Ramón	Costa Rica	San José	Pérez Zeledón
Mexico	Quintana Roo	Chetumal	Costa Rica	Guanacaste	Liberia
Mexico	Tamaulipas	Río Bravo	Costa Rica	Guanacaste	Santa Cruz
Mexico	Sonora	Hermosillo	Mexico	Jalisco	Guadalajara
Mexico	Michoacán de Ocampo	Lázaro Cárdenas	Mexico	Nuevo León	Linares
Mexico	Sinaloa	Ahome	Mexico	Nayarit	Tepic
Mexico	Tabasco	Comalcalco	Mexico	Oaxaca	San Juan Bautista Tuxtepec
Mexico	Michoacán de Ocampo	Uruapan	Mexico	Oaxaca	Tehuantepec
Mexico	Quintana Roo	Solidaridad	Mexico	Puebla	Puebla
Mexico	Sonora	Cajeme	Mexico	Nuevo León	Montemorelos
Mexico	Veracruz de Ignacio de la Llave	Boca del Río	Mexico	Nayarit	Santiago Ixcuintla
Mexico	Veracruz	Coatzacoalcos	Mexico	Nuevo Leon	Garza Garcia
Mexico	Quintana Roo	Felipe Carrillo Puerto	Mexico	Querétaro	Querétaro
Mexico	Sinaloa	Culiacán	Mexico	Nuevo León	Gral. Zuazua
Mexico	Tamaulipas	Ciudad Victoria	Mexico	Nuevo León	Juárez
Mexico	Tamaulipas	Nuevo Laredo	Mexico	Oaxaca	Juchitán de Zaragoza
Mexico	Sinaloa	Mazatlán	Mexico	Nuevo León	Cadereyta Jiménez
Mexico	Morelos	Jiutepec	Mexico	Nuevo León	Santa Catarina
Mexico	Jalisco	Zapopan	Mexico	Nayarit	Bahía de Banderas
Mexico	Sinaloa	Guasave	Mexico	Nuevo León	García
Mexico	Tabasco	Cardenas	Mexico	Puebla	Xicotepec
Mexico	Guerrero	Acapulco de Juarez	Mexico	Yucatan	Uman

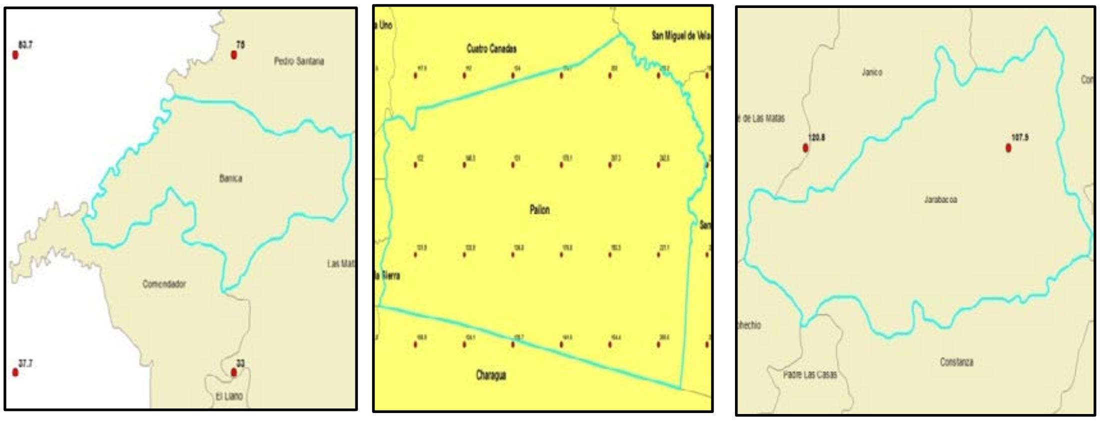
Source: Authors’ calculations.

Figure 1. Causal Pathways from Climate and Socioeconomic Conditions to Vulnerability and Health Outcomes



Source: Authors' compilation.

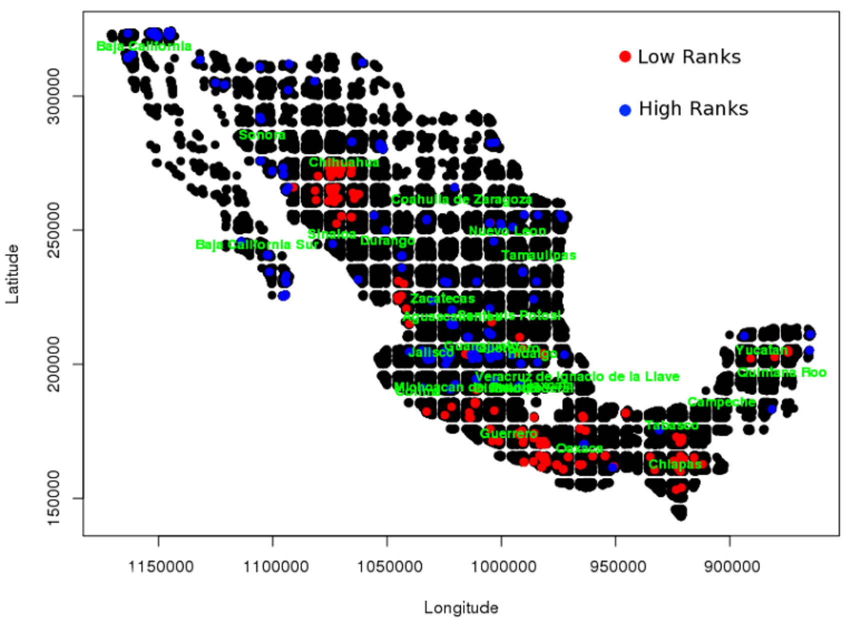
Figure 2. Cases Matching Geo-Referenced and Climate Data



Case 1. None-to-One **Case 2. Many-to-One** **Case 3. One-to-One**

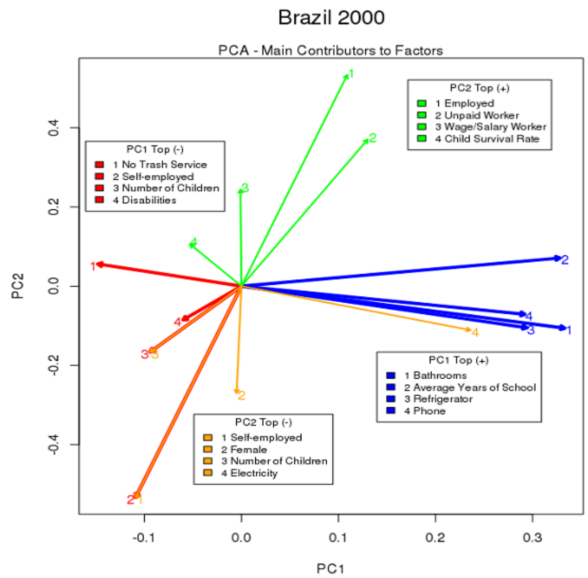
Source: Authors' compilation.

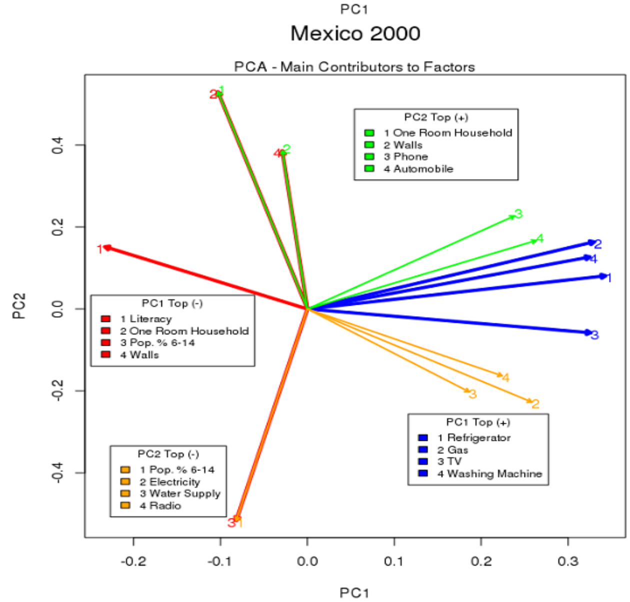
Figure 3. Sensitivity Index: Top and Bottom 50 Municipalities in Mexico (2000 Census).



Source: Authors' calculations.

Figure 4. Biplots: Principal Component Analysis, Brazil and Mexico 2000





Source: Authors' calculations.