Climate Change and Malaria in Brazil: A Research Proposal

Andrew Roberts
5/1/2018

Despite the widespread success of malaria control in Brazil, transmission rates remain high in the Amazon River Basin. Climate change poses a threat to continued malaria reduction, potentially influencing both behavioral and biological systems that affect the malaria burden. In this paper, I propose an empirical analysis to estimate the impact of climate change on Brazilian malaria incidence. In order to address obvious sources of endogeneity, I present three fixed effects models that seek to isolate the climate effect, assuming researchers have access to the necessary monthly hospital admissions and meteorological data relevant to this study.

1. Introduction

In everyday contemporary literature, the narrative concerning malaria in Brazil is primarily framed as a success story. In many respects, this characterization is certainly accurate; annual malaria cases have dropped from an estimated six million in the early 1940s to just over 124,000 in 2017.[1] Malaria reduction has been remarkable even over the course of the last decade, with the confirmed annual cases per 1000 people having dropped by a factor of about three from 2005 to 2016.[2] Despite this success, there exists a growing concern that continued climatic changes threaten to reverse some of the progress made in recent decades. A gap in epidemiologic research relating climate change and the incidence of vector-borne illness leaves the future burden of diseases such as malaria largely uncertain.[3] This paper introduces a framework in which to study the effects of changes in temperature, humidity, and rainfall on the incidence of malaria in Brazil. The data required to actually conduct the study proposed here does exist, but is not freely available to the public. Furthermore, the methodological framework set forth in this paper may be applicable in other study settings, such as in research dealing with different geographical areas and/or vector-borne diseases.

Brazil is an interesting country in which to study this problem because of its diverse climate and uneven distribution of malaria. While vector control measures have essentially eradicated malaria in most of Brazil, the disease burden continues to be a persisting problem in the Amazonian municipalities. In fact, the data has actually shown that morbidity has not been decreasing among at-risk populations (women and children) in these regions.[4] Changing climatic patterns clearly add further complications to this problem, but the mechanisms through which climate affects morbidity have little empirical backing. Any study striving to elucidate these mechanisms must account for the fact that changes in climate have the potential to alter the overall incidence of the disease, its geographical distribution, and human behavior.[5] Climatic variation may affect the survival and reproduction of mosquitos, the temporal pattern of vector activity (i.e., bite rates), and the rates of survival and reproduction of the pathogens living within the vectors.[6] Further complicating matters, in order to truly identify climate signals and their effect on malaria morbidity, a well-executed study must differentiate between long-term changes in climate norms (e.g., mean temperatures over the long-term), climate variability about these norms in a shorter timeframe, and isolated extreme events.[7]

In order to deal with the complexity of this research question, I propose a long-term longitudinal study with data aggregated at the municipality level. My proposed dataset consists of three primary explanatory variables that fall under the climate umbrella: temperature, rainfall, and humidity. This paper seeks to develop an empirical structure that would allow researchers to establish a causal relationship between these variables and malaria morbidity, filling the gap in existing literature.

If this study were implemented, I would expect to find positive relationships between the meteorological variables and malaria incidence. Based on the relevant literature, it seems likely that the precipitation effect would be the largest in magnitude. A large source of uncertainty stems from the fact that the estimated coefficients would also be detecting changes in human adaptation to the climate.

Untangling these contrasting effects is largely beyond the scope of this paper, but remains an important area for future research.

The remainder of this paper is organized as follows. The following section provides a theoretical background in climate econometrics and statistics. Section 3 describes the data required to carry out this study, as well as background information on malaria and climate in Brazil over the proposed study period. Section 4 discusses empirical strategies, while 5 outlines hypothesized results that may be obtained utilizing these strategies. Section 6 briefly examines some of the potential mechanisms underlying the climate-malaria relationship and proposes some basic methods to test their significance. Section 7 concludes.

2. Theoretical Background

In this section I seek to summarize some of the econometric advances in estimating climate effects on populations in order to provide the conceptual underpinnings of my proposed methodologies. First, it is important to specify a concrete, mathematical distinction between climate and weather that will underlie the econometric models presented later in this paper.

At any point in time *t* and location in space *i* one can imagine a vector of random variables characterizing atmospheric and oceanic conditions. In the context of this proposed study, this vector would look like:

$$\boldsymbol{v}_{it} = [temperature_{it}, precipitation_{it}, rainfall_{it}]$$

We imagine that $\boldsymbol{v}_{it}$ is drawn from a multivariate probability distribution defined over some time interval $\tau = [t_1,\ t_2]$.

$$\boldsymbol{v}_{it} \sim \varphi(\boldsymbol{C}_{i\tau})\ \forall t \in \tau$$

where $\boldsymbol{C}_{i\tau}$ is a random vector of K sufficient statistics characterizing the expected distribution of $\boldsymbol{v}_{it}$. For now, I simply consider the first moment of the distribution so $\boldsymbol{C}_{i\tau}$ can be thought of as $\boldsymbol{C}_{i\tau} = [\overline{temperature_{it}}, \overline{precipitation_{it}}, \overline{rainfall_{it}}]$. However, it is possible that other parameters (e.g., maximum temperature), better characterize this distribution.[i8]

Over the period of time τ it is unlikely that the observed values of the meteorological variables will exactly equal the expectations predicted by $\varphi(\boldsymbol{C}_{i\tau})$. Therefore, over the same interval of time, there exists an *empirical* distribution $\varphi(\boldsymbol{c}_{i\tau})$ that describes the realized distribution of the states in $v_{i,t\in\tau}$. The K parameters in $\boldsymbol{c}_{i\tau}$ are analogous to those in $\boldsymbol{C}_{i\tau}$ except that they characterize the *realized* distribution, rather than the *expected* distribution. Hence, for a specific period of time and location in space, I define climate as $\boldsymbol{C}_{i\tau}$ and weather as $\boldsymbol{c}_{i\tau}$. The interval of time, τ, over which climate is defined is somewhat arbitrary, as events such as the El Niño Southern Oscillation (ENSO) can alter the expected distribution of meteorological variables.

Now that this distinction has been made, the setting for the proposed analysis becomes much clearer. The ultimate goal is to estimate the average treatment effect of climate on the incidence of malaria in the Brazilian population. The impact of climate on this outcome variable can manifest itself in two ways: through direct biological and physical effects, and through human beliefs about the

[i] Mendelsohn (2016) notes that research has shown the second moment of climate distributions is important.

climate that may alter their behavior. Therefore, the outcome of interest can be written as

$$Y = Y(\boldsymbol{c}, \boldsymbol{b} \mid \boldsymbol{C})$$

where $\boldsymbol{b}$ is an N-dimensional vector comprising all actions affected by beliefs and both **c** and **b** are conditional on climate. Therefore, in estimating the impact of climate on malaria using the empirical methods discussed later, the model would pick up the combination of these two effects.[9]

An empirical model explaining the impact of climate on malaria incidence seeks to estimate

$$\beta = E[Y_{i\tau}|C_{i\tau} + \Delta C_{i\tau}, x_{i\tau}] - E[Y_{i\tau}|C_{i\tau}, x_{i\tau}]$$

where $x_{i\tau}$ is a vector of observable, non-climatic variables that affect Y and $\Delta C_{i\tau}$ is a change in climate.[10]

In response to the growing interest in linking climate change to economic outcomes, a great deal of research has been conducted in recent years exploring econometric techniques to estimate the effects described above. Deschenes and Greenstone (2011) noted the shortcomings of using cross-sectional data to study this problem, instead recommending the use of panels across time to avoid obvious sources of omitted variables bias.[11] The models outlined in section 4 follow from these findings. These models assume that a municipality *i* is comparable to itself over time—in other words, conditional on observables, that there exist no time varying unobserved factors that correlate with both explanatory variables and the outcome variable.[12] Furthermore, in order to approximate the average climate effect, the panel model makes a second implicit assumption,

$$E[Y_i|c_\tau] - E[Y_i|C_1] = E[Y_i|C_1 + (c_\tau - C_1)] - E[Y_i|C_1] = E[Y_i|C_1 + \Delta C] - E[Y_i|C_1]$$

that "the effect of a marginal change in the distribution of weather is the same as the effect of an analogous marginal change in the climate."[13]

3. Background and Data

To untangle the effects of climate change on malaria morbidity in Brazil, I propose a municipality-level panel data approach. This would entail collecting a longitudinal dataset composed of climate and health data from Brazil's 5,570 municipalities. In order to uncover trends over the medium and long-term, I suggest the time horizon span at least a decade; more concretely, in this paper I explore the potential of a hypothetical dataset covering the period 1990-2005, aggregated at the year-month level. This timeframe would conceivably offer sufficient variation in each of the relevant variables, as Brazil has experienced dramatic changes both in its climate and disease burden over the last few decades.

3.1 Climate

3.1.1 Historical Background

Encompassing nearly 50% of the landmass in South America and spanning latitudes from 5°N to 32°S, Brazil offers significant geographical variation in its climate.[14] Despite being contained primarily in the tropics, its unique topographical and geographical features such as the Amazon Basin in the north and the Brazilian Plateau in the south drive many of the disparities in temperature, rainfall, and humidity across the country. Parsing the long-term trends across these regions from the shorter-term variability and isolated weather events is necessary in any study hoping to reveal a link between climate change and disease. There already exists a growing body of research that has attempted to identify long-term climate signals in Brazil. The Met Office Hadley Center concluded in their 2011 report that a general period of warming has occurred in northern, eastern, and southern Brazil from 1960 to 2010.[15] Furthermore, an increase in winter temperatures averaged across the country has been identified in the same period.[16] In regards to rainfall, between 1960 and 2003 there has been a small increase in total annual precipitation over Brazil, but variation is largely attributable to inter-annual and decadal variability, rather than long-term trends in climate.[17] Other studies have documented a systematic increase of heavy precipitation events since the 1950s but have linked intra-seasonal variations in extreme rainfall events to El Niño and La Niña.[18] Historically, ENSO events have been associated with severe droughts in Amazonia, but the most recent have had relatively more mild impacts.[19] Over the course of the proposed study period, five El Niño and five La Niña events have been recorded, with substantial variation in their intensities. As a result of these events, in addition to other climatic factors, the late 1990s and early 2000s have marked a period of turbulent fluctuations in rainfall, the most notable being the severe Amazonian drought in 2005.[20]

3.1.2 Data

In order to properly conduct the study proposed in this paper, it would be necessary to obtain daily averages of temperature, rainfall, and humidity metrics for each municipality in Brazil over the period 1990-2005. While this dataset seems highly idealized, technological improvements in weather recording technology have opened the door to a vast collection of new data previously unavailable to researchers. Advances in satellite technology and extrapolation techniques have contributed to the growth of high-resolution gridded datasets that have been used frequently in the research community. A particularly relevant example is Xavier et al's (2015) work in developing .25° x .25° grids of daily temperature, precipitation, and relative humidity in Brazil over the period 1980-2013.[21]

3.2 Malaria

3.2.1 Historical Background

While a complete history of malaria transmission and control in Brazil would date back to at least the sixteenth century, in this section I offer only a partial summary of the recent history relevant to this research proposal.

In the 1940s, Brazil recorded 4-6 million malaria cases per year among a population of 45 million people, with more than 50% of these cases occurring outside of the Amazon.[22] The 1950s shift to national control efforts—making heavy use of DDT and chloroquine (CQ)—contributed to a significant reduction in malaria cases.[23] By 1976, these efforts essentially eradicated malaria outside of the Amazon.[24] Despite the success in Brazil's other regions, a new interest in Amazon colonization in the 1970s-1980s led to mass population movements and, with them, a resurgence of malaria cases in northern Brazil. Construction, mining, and agricultural projects were some of the primary contributors to this resurgence, resulting in an increase in malaria cases by a factor of 2.4 between 1980 and 1985.[25] By 1986, 99% of Brazilian malaria cases occurred in the Amazon.[26]

The national government responded to this persistent problem with the Amazon Basin Malaria Control Project (PCMAM). Utilizing a strategy of ultra-low volume localized insecticide spraying, PCMAM managed to significantly reduce malaria incidence between 1989-1996.[27] However, depleted funding led to another resurgence between 1989-1999. In 1999, cases were recorded in every Amazonian state, with the Brazilian Amazon accounting for 99.7% of all malaria cases in the country.[28] In response, a subsequent Amazonian malaria control program, focusing on indoor insecticide spraying, drainage projects, and evaluation of new settlements, managed to reduce the number of malaria cases back to pre-resurgence levels (although success varied significantly by state).[29]

By 2005, malaria control passed to the National Program for Malaria Prevention and Control (PNCM). Around the time of its inception, 47 Amazonian municipalities accounted for 70% of malaria cases. A focus on early diagnosis and rapid treatment resulted in a decline in incidence between 2010-2013.[30] Today, the Amazon Basin continues to register about 99% of all Brazilian malaria cases, with regional spikes often associated with large-scale public works projects or economic shifts.[31]

3.2.2 Data

The primary dependent variable of interest in this study is the monthly incidence of malaria aggregated at the municipality level. Malaria mortality rates have been greatly reduced in Brazil due to improvements in surveillance and treatment so a measure of morbidity is preferred in order to more fully capture the burden of the disease. I propose using monthly hospital admissions, a commonly used measure in epidemiologic research, as a proxy for malaria incidence. It should be noted that this would almost certainly be an underestimate of the true burden of the disease.

Although Brazil's health ministry does not issue frequent epidemiological notices on the incidence of malaria, there exists a significant body of municipality-level data not freely available to the public.[32] Barring some reporting delays in the Brazilian countryside, the Epidemiological Surveillance Information System for Malaria (SIVEP-Malaria) gathers data in real-time, which may be accessed online with a login.[33]

4. Empirical Method

This section outlines the primary proposed empirical method to establish a causal link between malaria incidence and climate. Secondary analyses and robustness tests are presented in the following sections. The relationships between malaria and the meteorological variables in this analysis are expected to be non-linear, as noted frequently in the relevant literature.[34] Furthermore, of particular interest are not only the individual marginal effects of the three primary explanatory variables, but also their associated interactions. A non-parametric, binning approach such as that used by Burgess et al (2017) allows for considerable flexibility in the response function, but at the expense of the parsimony of the model.[ii] [35] However, this model does allow the explanatory variables to capture the distribution of daily temperature and relative humidity, which has been shown to be ideal in these types of models.[36] A more parameterized model loses the daily variation but remains fairly parsimonious, even with interaction terms and lags included. Because of these

[ii] If each of the three explanatory variables have 7 bins then the total number of terms would be 147 (excluding lags/control variables and assuming interactions only between bins of different variables)

tradeoffs, I present two separate models that may be appropriate in this empirical setting.

4.1 Binning approach

This first method is largely analogous to the model used by Burgess et al to estimate the impact of temperature on mortality in India.[37] This model is highly flexible and captures the daily distribution of the meteorological variables but is not able to accommodate interaction terms in the interest of parsimony. Estimates are obtained from the following model:

$$Y_{it} = \sum_{j=1}^{8} \theta_j TMEAN_{itj} + \sum_{k=1}^{8} \delta_k 1\{RAIN_{it}\ in\ octile\ k\} + \sum_{l=1}^{8} \beta_l HMEAN_{itl} + \lambda X_{it} + \alpha_i + \gamma_t + u_{it} \qquad (1)$$

where Y_{it} is the log of hospital admissions in municipality *i* in year-month *t*. The TMEAN and HMEAN bins capture the daily distributions of temperature and relative humidity, respectively. Both are constructed identically and require a gridded daily dataset of temperature and humidity. The variable $TMEAN_{itj}$ indicates the number of days in municipality *i* and year-month *t* on which the daily mean temperature fell into temperature bin *j*. The construction is identical for relative humidity. The decision to include eight bins for each of these categorical variables is somewhat arbitrary, but follows from the work done by Burgess et al.[38] Due to data limitations, the first and last bins characterize the temperature ranges <65° F and >95° F, respectively, while the remaining six bins are broken into bins 5° F wide. This formulation implicitly assumes the impacts of interest only depend on daily means, that the impacts are constant within 5° F intervals, and that the sequence of varying weather patterns is unimportant.[39] Conditional on the date range of the study and the humidity distribution itself, the binning approach will be similar for the humidity bins. The optimal number of bins for temperature and relative humidity should be chosen empirically and will almost certainly vary depending on data limitations. I consider eight bins for both in this proposed model simply to maintain consistency with recent literature.

The precipitation terms on the right-hand side of the equation also follow from the empirical work conducted by Burgess et al (2017).[40] The variable $1\{RAIN_{it} in\ octile\ k\}$ is an indicator that returns one if the cumulative rainfall in municipality *i* and year-month *t* falls into the k^{th} octile of the distribution of rainfall in municipality *i* over the time horizon of the dataset; otherwise, it returns zero. Once again, choosing the number of bins is an empirical question that should be decided conditional on the dataset. Alternatively, one could model rainfall using a binning approach based on daily means if the available data is sufficient.

The vector *X* contains covariates that could plausibly be correlated with both climate and malaria incidence. Potential sources of endogeneity are discussed in section 4.3.

This specification also includes municipality fixed effects, α_i, which remove time-invariant observed and unobserved characteristics specific to municipalities. The time fixed effect, γ_t, is shorthand for year-month dummy variables controlling for municipality-constant characteristics that vary across time. Correlation among the idiosyncratic errors, u_{it}, is allowed within municipalities, meaning standard errors are clustered at the municipality level.

Overall, this specification has some appealing characteristics. Its non-parametric form allows for a very flexible response function. It also aligns with recent findings in epidemiologic literature that the impact of climate on malaria transmission is dependent on *daily* temperature variation.[41] The time-series aspect of this model accounts for cross-municipality differences, removing a large potential source of omitted variables bias. However, this specification is still susceptible to omitted variables bias in that unobserved, time-varying characteristics that are correlated with both malaria hospital admissions and any of the explanatory variables (after conditioning on the time dummies and controls) will bias the coefficients.[42] Furthermore, in order to reasonably estimate this model, interaction terms and lags have been omitted, leaving out a potentially significant source of explanatory power.

4.2 Parameterized approach

In order to construct a more parsimonious model that allows for the addition of interactions and lags, I present a second model that abandons the binning strategy in favor of a more parameterized specification. Estimates using this alternative strategy are found using the following model:

$$Y_{it} = \theta_1 T_{it} + \theta_2 T_{it}^2 + \delta_1 P_{it} + \delta_2 P_{it}^2 + \beta_1 H_{it} + \beta_2 H_{it}^2 + \eta Z_{it} + \lambda X_{it} + \alpha_i + \gamma_t + u_{it} \quad (2)$$

where the outcome variable, vector of controls, fixed effects terms, and idiosyncratic error term are specified identically to model (1). The variables T_{it}, P_{it}, and H_{it} represent the respective averages of temperature, precipitation, and rainfall in municipality *i* and year-month *t*. In order to capture the non-linearity of the relationships, a second-order polynomial term is included for each of the three primary explanatory variables. This model implicitly assumes that the impact on the dependent variable is governed exclusively by the monthly averages of the variables, potentially missing out on the explanatory power of the daily distribution. However, the more succinct parameterized form of this specification allows for the addition of interaction terms, which are included between every possible set of two distinct meteorological variables and are written succinctly in the vector Z.[iii] For the moment, this design only measures the contemporaneous impact of climate on malaria incidence. To detect the effect climatic changes may have on malaria in future months, the following model adds one-month lags for each of the climatic variables:

$$Y_{it} = \phi_{it} + \kappa_1 T_{i,t-1} + \kappa_2 T_{i,t-1}^2 + \omega_1 P_{i,t-1} + \omega_2 P_{i,t-1}^2 + \rho_1 H_{i,t-1} + \rho_2 H_{i,t-1}^2 \quad (3)$$

where ϕ_{it} denotes the right-hand side of equation (2). Currently, this model assumes that the lagged effect of one variable is constant across all values of the other variables (whereas this is not true of the contemporaneous effects). However, additional dynamics could easily be added to this baseline framework. Although now containing interaction effects and more fully specified dynamics, this fixed effects model is vulnerable to the same forms of omitted variables bias as models (1) and (2). Opting for a parametric approach that loses out on daily variation but incorporates additional dynamics, this specification offers a distinct framework in which to explain the relationship between climate and malaria.

[iii] If all possible interactions are included, $Z \in \mathbb{R}^{11}$

4.3 Endogeneity concerns

The primary concern shared by both of these models is the presence of unobserved factors affecting malaria transmission that are also correlated with climate. While the fixed effects address the concerns related to cross-sectional endogeneity, time-varying factors still have the potential to bias regression coefficients.

Perhaps the most important covariate that should be included in the vector **X** is municipality population. As discussed in the background section, population movements have had significant effects on regional malaria transmission; including a variable representing population would plausibly control for most of the variation related to human migration, though it is insufficient if malaria incidence is a function of intra-municipality population movements (which could be driven by climate). Considering the high degree of population movement in Brazil over the last couple decades, this variable would plausibly contain enough temporal variation to be estimated in a fixed effects framework.

A second concern is that malaria control measures (e.g. insecticide spraying) may be correlated with climate. For example, increased temperatures may increase malaria transmission and induce spending on malaria reduction measures. Assuming these measures are actually effective, in this scenario the estimated coefficients will understate the effect of temperature on malaria incidence. In order to address this concern, one potential avenue is to introduce a proxy that quantifies the "effort" put into malaria reduction in each municipality. A simple approach would be to control for monthly government spending in each municipality. This approach neglects personal spending habits related to malaria; therefore, assuming no other endogeneity issues, the coefficients will estimate the climate effects on malaria caused by both individual behavioral changes and biological impacts.

Another concern arises from potential measurement error in the outcome variable. The primary outcome of interest (malaria-related hospital admissions) almost surely represents an underestimate of true malaria incidence. If within-municipality underreporting is time invariant, then the measurement error poses no problem to the analysis. However, if reporting accuracy changes over time and is correlated with climate, it could bias the estimated coefficients. For example, if malaria reporting, on average, improves over time (i.e., becomes less of an underestimate) as average temperatures increase, the estimated temperature coefficients will be biased upwards, partly reflecting reporting improvements that have nothing to do with temperature. To address this concern, researchers may opt to utilize *estimated*, rather than *reported*, malaria cases as the outcome variable.[43] However, there is no guarantee that this adjusted metric completely eliminates measurement error.

4.5 Robustness checks

The empirical strategies described in this section serve as a suitable baseline analysis, motivating the empirical setting in which to estimate the climate's effect on malaria. In this section I further the analysis by suggesting some basic techniques to investigate the robustness of model results.

The timing of the biological and behavioral responses induced by climate change is clearly important in estimating the impact on malaria. Model (3) accounts for this timing by introducing one and two-month lags, but sacrifices some flexibility by imposing a parametric form on the relationship between climate and malaria. I suggest an alternate specification that models climate non-parametrically, but only includes two bins for each climate variable. These bins should be chosen such that moving from one bin to another represents crossing a key threshold relevant to malaria incidence (this can be explored empirically and also be based on existing biological evidence). As a more parsimonious non-parametric model, it can feasibly accommodate a small number of lagged terms for temperature, rainfall, and humidity. Therefore, this model can be used to demonstrate whether coefficients change appreciably when the number of lags is changed or if lags are removed altogether. Ideally, the interpretation of the coefficients will agree with the results found using the other models. For an example in which a simplified model is used to check for the robustness of coefficient estimates to the addition of lags, see Burgess et al (2017).[44]

The second robustness check I propose is a simple lead falsification test, in which the climate variables one-month in the future are used as explanatory variables for the current month. By verifying that the coefficients on these leads are small in magnitude and statistically insignificant this check acts as a basic sanity check.

Next, I suggest estimating more conservative standard errors to investigate the validity of statistical inference conducted using these models. Many Brazilian health statistics are collected by federative district, so I propose an alternate specification in which standard errors are clustered by federative district instead of municipality. This approach will produce larger standard errors and thus lead to more conservative inference. One can then determine the robustness of results obtained from statistical inference when standard errors are clustered at different levels of aggregation.

The final set of robustness checks I propose addresses concerns related to the outcome variable. To investigate the potential effect of measurement error in the analysis, researchers can obtain malaria datasets collected by different organizations and determine whether the use of different data appreciably affects the results. Additionally, as noted previously, one can substitute *estimated* for *reported* malaria-related hospital admissions in an effort to minimize the measurement error in the dependent variable.

4.6 A note on implementation

It should be noted that these proposed models are intended to act as general guidelines for empirical analysis. The final specification in a real research setting will obviously depend on the data and should undergo rigorous empirical tests. I present the models in the above order because I envision the earlier models informing decisions to me made in the later ones. In other words, the parameterizations of models (2) and (3) may be based on results found by estimating model (1). For example, model (1) may reveal that maximum temperatures govern the relationship between temperature and malaria incidence better than averages. In this case, maximum temperatures should replace mean temperatures in the parameterized models.

5. Expected Results

If this study were implemented, I would expect to find positive relationships between the meteorological variables and malaria incidence. Because of the documented importance of the presence of standing water in malaria transmission, I hypothesize the precipitation coefficient to be the largest in magnitude. Conditional on rainfall and temperature, it seems likely that the estimated humidity effect would be relatively smaller than that of the other climate variables. However, the extent to which behavioral adaptation may alter the estimated coefficients is largely uncertain. I would expect self-protection expenditures to somewhat lower these coefficients, but not so much as to drive them into negative territory. A more in depth discussion of these behavioral mechanisms in provided in the following section.

6. Mechanisms

I now turn to a brief discussion of secondary analyses that could be employed to further understand the mechanisms underlying the relationship between climate and malaria. As stated before, the estimates obtained using the methods outlined in the previous section combine the behavioral and biological effects induced by a change in climate. This section aims to provide general approaches that would help to untangle these effects. It should be noted that the empirical tools and data discussed above are likely not sufficient to accurately estimate long-run impacts such as population movements (both human and vector). Therefore, this section focuses on shorter-term changes that may help explain the model results.

Within the short-run there are various factors that influence whether the regression estimates should be interpreted as upper or lower bounds. An obvious impact supporting the lower bound interpretation is that climate change may be increasing human adaptation. This would imply that individuals are shielding themselves from the increased risk of contracting malaria by purchasing more insecticide-treated bed nets or avoiding areas known to have dense mosquito populations. On the other hand, if changes in climate induce behavioral effects such as increasing individuals' tendencies to leave windows open or to go outside more, the true climate effect on malaria may be smaller than the estimated effect. In general, it is easy to tell contrasting stories describing the behavioral impact as either amplifying or dampening the true biological effect.

Some of these effects may be tested empirically, while others are likely not feasible to study due to data constraints. For example, the necessary data to estimate the effect of climate change on the probability of an individual leaving their window open at night is likely not obtainable. The most reasonable avenue for further research is to estimate the impact of climate on self-protection expenditures. This data could feasibly be obtained from Brazilian clinics that provide insecticide-treated bed nets or comparable malaria avoidance products. The resulting regression specification would be identical to those described in section 4, except that malaria-avoidance expenditures would replace malaria incidence as the dependent variable.

7. Conclusion

As climate change continues to alter Earth's atmospheric and oceanic conditions, the resulting behavioral and biological changes threaten to cloud the future trajectory of many of the vector-borne diseases still present in parts of the world today. A fundamental understanding of the effect a changing climate will have on these diseases could be instrumental in the development of new programs and interventions aiming to mitigate the impacts of vector-borne disease. Currently, there is a significant lack of evidence on this subject; this paper suggests an empirical framework that could potentially be utilized to begin to fill this gap in existing research.

The primary contribution of this paper is a proposed model that makes an initial attempt to estimate the impact of climate change on malaria incidence in Brazil. While unable to detect long-term climate effects, this model would theoretically be capable of picking up on impacts occurring on a short to medium timeframe. The main expected result is a positive association between increases in temperature, humidity, and rainfall and malaria incidence. However, the extent to which adaptation would play a role in the estimation is unclear. Elucidating the mechanisms underlying the climate-malaria relationship is difficult considering current data availability. This paper makes some basic suggestions but this question largely remains an open area for future research.

Generally, the methods described here make an initial attempt at uncovering a causal relationship between climate change and malaria. The models are by no means completely refined, but may serve to guide future research on the subject.

This paper has underscored the importance of continued investigation into this topic, as a better understanding of the link between climate change and malaria could yield numerous public health benefits in the future.

References

[1] "Malaria Country Profile," *World Health Organization*, n.d.
[2] "Malaria Country Profile."
[3] R.S. Kovats et al., "Climate Variability and Change in the United States: Potential Impacts on Vector- and Rodent-Borne Diseases," *The Royal Society* 356 (2001): 1057–68.
[4] Oliveira-Ferreira et al.
[5] Kovats et al., "Climate Variability and Change in the United States: Potential Impacts on Vector- and Rodent-Borne Diseases."
[6] Kovats et al.
[7] Kovats et al.
[8] Robert Mendelsohn, "Measuring Weather Impacts Using Panel Data," *Journal of Economic Literature* 51 (n.d.).
[9] Solomon M. Hsiang, "CLIMATE ECONOMETRICS," *National Bureau of Economic Research* 22181 (n.d.).
[10] Hsiang.
[11] Olivier Deschênes and Michael Greenstone, "CLIMATE CHANGE, MORTALITY, AND ADAPTATION: EVIDENCE FROM ANNUAL FLUCTUATIONS IN WEATHER IN THE US," *National Bureau of Economic Research* 13178 (June 2007).
[12] Hsiang, "CLIMATE ECONOMETRICS."
[13] Hsiang.
[14] Met Office, "Climate: Observations, Projections and Impacts: Brazil," *Department of Energy & Climate Change*, n.d.
[15] Met Office.
[16] Met Office.
[17] Met Office.
[18] Jose A. Marengo, "Climate Change and Extreme Events in Brazil," *Brazilian Foundation for Sustainable Development*, n.d.
[19] Marengo.
[20] Met Office, "Climate: Observations, Projections and Impacts: Brazil."
[21] Alexandre C. Xavier, Carey W. King, and Bridget R. Scanlon, "Daily Gridded Meteorological Variables in Brazil (1980–2013)," *Royal Meteorological Society* 36 (October 8, 2015): 2644–59.
[22] Sean Michael Griffing et al., "A Historical Perspective on Malaria Control in Brazil," *Brazilian Malaria and Malaria Control* 110, no. 6 (July 14, 2015): 701–18.

[23] Griffing et al.
[24] Griffing et al.
[25] Griffing et al.
[26] Griffing et al.
[27] Griffing et al.
[28] Griffing et al.
[29] Griffing et al.
[30] Griffing et al.
[31] Griffing et al.
[32] Judith Recht et al., "Malaria in Brazil, Columbia, Peru and Venezuela: Current Challenges in Malaria Control and Elimination," *Malaria Journal* 16, no. 273 (2017).
[33] Recht et al.
[34] Tamma A. Carleton and Solomon M. Hsiang, "Social and Economic Impacts of Climate," *Sciencemag.Org* 353, no. 6304 (September 9, 2016): 1112–30.
[35] Robin Burgess et al., "Weather, Climate Change and Death in India," *National Bureau of Economic Research*, April 20, 2017.
[36] Krijn Paaijmans P. et al., "Influence of Climate on Malaria Transmission Depends on Daily Temperature Variation," *PNAS* 107, no. 34 (August 24, 2010): 15135–39.
[37] Burgess et al., "Weather, Climate Change and Death in India."
[38] Burgess et al.
[39] Burgess et al.
[40] Burgess et al.
[41] Krijn Paaijmans P. et al., "Influence of Climate on Malaria Transmission Depends on Daily Temperature Variation."
[42] Hsiang, "CLIMATE ECONOMETRICS."
[43] "Malaria Country Profile."
[44] Burgess et al., "Weather, Climate Change and Death in India."