APPROVED
By Sebastian Bauhoff at 11:57 am, Nov 20, 2017

Note: this pre-analysis plan is Phase 2 of two phases of this project. In Phase 1 we constructed and ran the code for a single country, Liberia. In this pre-analysis plan for Phase 2 we describe the analysis that we will perform on a full set of countries. Deviations from the original pre-analysis plan are marked [D] below.

Scope

1. Sample selection and geographic scope of analysis
    a. We first identified all Demographic and Health Surveys (DHS) that asked about fever in children, including the subset of surveys that included tests for malaria.
    b. We limited the study to surveys fielded between 2001 and 2014 (the period for which temperature and precipitation data is available)
    c. We drop those 2014 surveys where some observations were gathered in 2015 for which we do not have temperature data.
    d. We choose to focus only on South and Southeast Asia, Sub-Saharan Africa, and Latin America/Caribbean.
    e. Within those surveys, we will focus only on rural areas, as identified by the DHS data, and omit urban areasThe final list of countries and the applicable sample sizes are listed in Table 1 below.

Data

2. Data sources – dependent and independent variables

    a. Primary dependent variable: Malaria in children under age 5 at the time of the survey (Source: DHS). DHS employs two malaria blood tests: rapid test and microscopy. We will use the rapid test as the primary dependent variable, as it was used in more countries than the microscopy test. We will follow the definitions used by the DHS.
    b. Alternative dependent variable I: microscopy test.
    c. Alternative dependent variable II: Fever in children under age 5 in the two weeks preceding the survey (Source: DHS). We will follow the definitions used by the DHS.
    d. We will run all tests on all three dependent variables. For all dependent variables, we use the maximum available sample. That is, we use all observations with fever data, even if many of them do not also have rapid and microscopy measures. Similarly, we use all observations that have rapid tests even if some of them do not also have microscopy outcomes.
        i. As a sensitivity analysis, we implement the main specification on a restricted sample of observations that have all three outcome variables, for each of the three outcomes. That will indicate whether differences in estimated effects are due to differences in the samples. [D]
        ii. A fraction of surveyed children under five are noted as having died prior to the survey. We remove these observations from the sample. To the extent that some of these deaths were due to malaria (malaria is a leading cause of child mortality), malaria rates in the sample may be underestimated, and the relationship between deforestation and malaria cases may be underestimated. As a robustness check, we perform an analysis with death as the dependent variable, with the ex ante hypothesis that some substantial fraction of child deaths are due to malaria and thus child deaths would follow the same pattern as malaria prevalence, albeit with

the possibility for noise in cause of death to diminish the significance of the effect of forest variables. [D]

e. Independent variables (treatment)
    i. Deforestation during the year of the survey, as test of "land-cover change" hypothesis (Source: updated version of Busch and Engelmann 2015, who classified annual 30 m Landsat-derived tree-cover loss data (Hansen et al 2013/GFW) into forest or non-forest using a tree-cover threshold of 25% for 2001-2012, using same methods but adding 2013-2014).
    ii. Forest cover, as test of "land-cover" hypothesis (Source: Busch and Engelmann 2015, who classified 30 m Landsat-derived tree-cover data (Hansen et al 2013) into forest or non-forest using a tree-cover threshold of 25%). Forest cover during the year of the survey is inferred by subtracting previous years' forest loss from forest cover in year 2000.
f. Weights. We will use the sampling weights supplied by the DHS (see below for details).
g. We discard observations with data missing from at least one field, so that summary statistics accord with the sample used in multivariate analyses.
    i. In the case of dependent variables that were not collected for entire surveys; e.g. malaria rapid test or malaria lab test, we do not drop the observations from the sample.

3. Summary statistics:
    a. Present table of summary stats of dependent variables, independent variables, included control variables, pooled across all surveys, listing all variables.
    b. Survey-by-survey, listing country, year surveyed, sample size for malaria, sample size for fever
    c. Map of survey clusters across the world
    d. Timeline of DHS field work

4. Superficial analyses
    a. Testing for a first-order correlation between the key outcome variables (malaria rapid test; malaria microscopy; fever) with each other and with the main independent variable (deforestation)
        i. For each of the nine pairs we report correlation coefficient and significance (p-value)
    b. Heatmaps, to visually inspect for consistent relationship between forest cover/forest cover change and malaria:
        i. Malaria rapid test % (color) vs forest cover (x-axis, 25 increments from 0-100%) and deforestation (y-axis, 25 increments from 0 to 90th percentile)
        ii. Malaria microscopy % (color) vs forest cover (x-axis, 25 increments from 0-100%) and deforestation (y-axis, 25 increments from 0 to 90th percentile), for all DHS surveys that have malaria tests
        iii. Fever % (color) vs forest cover (x-axis, 25 increments from 0-100%) and deforestation (y-axes, 25 increments from 0 to 90th percentile), for all DHS surveys that have fever

5. Control variables
    a. There are limitations to inferences that can be drawn from first-order correlations and heatmaps.

b. We are concerned that third-factors could be correlated with higher or lower deforestation AND have an effect on malaria. Thus, we want to control for observable factors that have an effect on malaria.
c. But, what to control for? DHS reports hundreds of questions from which to draw possible covariates, and then there are other data sets too. So there are a nearly infinite number of possible permutations of variables to include, and if we ran enough tests we could likely find at least some showing a positive or negative association, even in the absence of a "true" relationship (Olken, 2015).
d. We choose variables for which there is a strong theoretic basis in the literature to have a *direct* effect on higher or lower malaria. We do not include variables that might have an *indirect* effect. Additionally, we seek to use variables that are universally available across surveys, and have a uniform interpretation across countries and time periods
e. We specify in a pre-analysis plan which control variables we will include:
    i. Theorized proximate causes of malaria assumed to be unaffected by deforestation:
        a. Temperature (in Celsius) during the month of the survey. Source: Terrestrial Air Temperature: 1900-2014 Gridded Monthly Time Series (Version 4.01). We use cardinal values and cardinal values squared, following inverted-U-shaped relationship from literature (Beck-Johnson et al 2013, Mordecai et al 2013)
        b. Precipitation (in mm) during the month of the survey. Source: Terrestrial Air Temperature: 1900-2014 Gridded Monthly Time Series (Version 4.01). We use cardinal values and cardinal values squared, following inverted-U-shaped relationship from literature (Parham and Michel 2010)
        c. Child age – use dummy variables for each year relative to Age<1
    ii. Theorized proximate causes of malaria potentially correlated with socioeconomic conditions associated with deforestation
        a. Housing quality, as proxied by a binary indicator of floor type (type (0 for unfinished eg mud, dirt; 1 for wood, cement, etc) [D]
            a. As a sensitivity analysis, we use a housing quality, index constructed by summing floor type (0 for unfinished eg mud, dirt; 1 for wood, cement, etc), and wall type (0 for unfinished eg mud; 1 for finished eg wood, cement), and roof type (0 for unfinished eg thatch; 1 for finished=metal, tile, etc;). Because wall type and roof type were missing from more than 20 survey waves we do not use this as the variable in the main analysis.
            b. See Table 2 below for codes of individual floor, wall, and roof type
        b. Exposure to standing water: Main water source (0 if open vs. 1 if pumped/piped)
        c. Access to health services: We use as a proxy variable a binary indicator of 1 if the child was delivered in a facility and 0 otherwise. We consider this a good indicator because birth has been universally experienced by children under 5 and birth in a facility is assumed to universally indicate better access to health services.
            a. Alternative proxy indicators such as 'child has received other vaccinations', are potentially less useful because recommended vaccines vary by country, some children may be too young to

have received vaccines, and there is probably a selection effect to vaccination campaigns, meaning that having been vaccinated might either mean better access to health services OR higher health risk.

b. However, only 85 surveys out of 97 asked about place of delivery. There does not appear to be another acceptable indicator of access to health care that is universally asked.
c. Rather than reduce the size of our sample, we will include access to health services only in a sensitivity analysis and in the mediation analysis (explained below)

iii. Avoiding behavior.
    a. In a sensitivity analysis, we add bed net – a binary variable that is 1 if "all children" or "some children" slept under a bed net last night; 0 otherwise. Because this bed net question was missing from more than 30 survey waves we do not use this as the variable in the main analysis. [D]
    b. Note: we are estimating the effect of deforestation at current levels of avoiding behavior. To the extent that deforestation increases malaria which increases avoiding behavior which dampens the increase in malaria, our estimates of the effect of deforestation on malaria are too small—that is, they have subtracted the effect of increased avoiding behavior component

f. We considered but chose not to include the following variables because they are not proximate, i.e. *direct*, causes of malaria:
    i. Wealth
    ii. Education
    iii. Remoteness
    iv. Concurrent forest-protection policy, ie. % of cell that is designated as protected

Hypothesis testing

6. Main multi-variate analysis (pooled cross-section with child-level observations)
    i. What specification to use? As with included variables, we pre-specify in order to avoid the potential to p-hack by applying log terms, fixed effects, functional forms, etc.
        a. Unit of observation:
            a. survey result for child under 5
        b. Functional form:
            a. Our preferred primary function form is logit, since our values are 0/1
            b. Additionally, we use linear probability models (OLS) as a robustness check.
        c. Modification of variables, if any
            a. All variables are entered as cardinal values, that is, implying a logistic relationship with the probability of malaria, with the exception of:

i. Forest cover, which is entered as forest cover+forest cover squared to test for hypothesized highest effects at intermediate values
ii. Temperature and precipitation, which use cardinal values and cardinal values squared, following inverted-U-shaped relationship from literature (Beck-Johnson et al 2013, Mordecai et al 2013; Parham and Michael 2010)
iii. Child age, which is entered in age classes

d. We cluster standard errors at the level of the grid cell because the exposure (forest cover; forest cover change) is common to all children in a grid cell.
    a. We will cluster on the DHS' primary sampling unit, as a robustness check.
e. We include survey fixed effects
f. Weighting [D]
    a. We weight all observations by the product of:
        i. within-survey weights provided by DHS, which is equal to the inverse likelihood of sampling to obtain a nationally representative sample
        ii. and, the relevant population of the country during the year of the survey (female population aged 15-49, drawn from the World Population Prospects: The 2017 Revision, File POP/7-3 "Female population by five-year age group, region, subregion and country, 1950-2100 "
        iii. In order to obtain a weighted sample that is representative of the aggregate population of all country-years surveyed
    b. As sensitivity analyses, we:
        i. Weight only on the within-country weights provided by DHS
        ii. Run the regression on an unweighted sample

7. Secondary multi-variate analysis: Imbalanced panel with cell-level observations
    a. In a subset of surveys we are fortunate enough to have multiple surveys. If it turns out that there are repeat measurements across years (i.e., subsequent DHS visited the same enumeration areas or at least areas within the same cell), this would enable a panel regression using cell-specific fixed effects. This lets us control for the effects of spatially variant, time-invariant *unobservables* as well as *observables.* That is, we can test whether a change in deforestation leads to a change in malaria, within a cell.
        a. Aside from the addition of cell-level fixed effects, we also add year dummy variables, using the year of the survey wave. All other features of the specification are the same as above
        b. Note: in the panel analysis, we can only test the land cover change hypothesis (increased deforestation has increased malaria) not the land cover hypothesis (intermediate levels of forest cover have highest malaria)

c. In the Liberia-only data the panel regression did not converge. However, we attempt it here again. With many more observations and the same number of variables, it is possible it will converge in the full sample.

8. Supplementary ex ante hypotheses

    a. Disaggregations
        i. Geographic variation: Disease ecology might be very different in different places – our ex ante hypothesis based on the literature is that deforestation in African and Latin American countries will have a larger effect on malaria than deforestation in Asian countries.
            a. We will run the specification for the pooled model but separately survey-by-survey, and report how many surveys show positive/negative/no significant correlation between malaria and forest cover change, and malaria and the joint forest cover+forest cover squared, in each region
            b. We will test for differences in effect between Africa and Latin America vs. Asia by adding interaction terms to the pooled model: Asia*deforestation; Asia*forest cover; Asia*forest cover squared. Our ex ante hypothesis is that Asia*deforestation will be negative and significant. We conduct this analysis only for fever, as the microscopy and rapid tests are only available in the African surveys.
        ii. Earlier vs later forest transition: We will add as an independent variable the interaction term deforestation*forest cover. Our ex ante hypothesis is that this term will be positive, ie. deforestation at higher forest cover will have a larger effect on malaria than at lower forest cover
        iii. Smaller vs larger cuts. We will add deforestation-squared as an independent variable. Our ex ante hypothesis is that deforestation-squared will be negative, meaning that the marginal effect of a hectare of deforestation on malaria will diminish as cuts increase in size
    b. Test of mediating effects
        i. Are housing quality, health care access, and water source mediating factors through which the effect of deforestation operates? We test for this using the methods of Keele et al 2015.
            a. First, we regress the housing quality index as a dependent variable on deforestation as the independent variable, including the other control variables from the standard model (temperature, precipitation, child age, bed net usage)
            b. We also regress health care access, as proxied for with children born in a clinic as a dependent variable, on deforestation as the independent variable, including the other control variables (temperature, precipitation, child age, bed net usage)
            c. We also regress water source as a dependent variable, on deforestation as the independent variable, including the other control variables (temperature, precipitation, child age, bed net usage)
            d. Then, for each of the three mediating factors, we calculate their predicted value:

a. with observed levels of deforestation, and
b. without deforestation (counterfactual scenario: deforestation=0)

e. Then, using the predictive model of malaria prevalence, calculate two predicted levels of malaria prevalence:
    a. with observed levels of deforestation and observed levels of housing quality, health care access, and water source
    b. with observed levels of deforestation but with the levels of housing quality, health care access, and water source predicted in the counterfactual scenario of no deforestation
f. The difference between the two predicted levels in d is the effect of deforestation on malaria via housing, health care, and water source.
g. We divide the level in f by the total effect of deforestation on malaria to calculate % of the effect of deforestation on malaria attributable to those three mediating factors

ii. Does land-cover effect mediate the effect of deforestation? We can't test for this using these methods because it violates the "sequential ignorability" assumption (Imai, Keele, and Yamomoto 2010) if land cover influences deforestation, as it likely does per Busch and Engelmann 2015. So we will not try to isolate the effect of deforestation via land cover.

iii. Does the specification of forest cover and deforestation influence the results?
    a. We will run the core regression with only deforestation, and not forest cover or forest cover squared
    b. We will run the core regression with only deforestation and forest cover, and not forest cover squared
    c. We will run the core regression with only forest cover and forest cover squared, and not deforestation

c. Temporal effects:
    i. We supplement deforestation in the concurrent year with three additional terms: deforestation 1-3 years ago; deforestation 4-6 years ago; and deforestation 7-9 years ago. Our ex ante hypothesis is that the first three terms will be significant, but deforestation 7-9 years ago will not be, following Singer and de Castro suggestion that 'frontier malaria' effect lasts 6-8 years. We run this test only for the subset of surveys conducted 2010 or later.
    ii. Lagged malaria incidence [D]: We run the analysis only for children age 0, as these children could only have acquired malaria during the concurrent period rather than a previous period
d. Other disaggregations and sensitivities are reserved for exploratory analysis to generate hypotheses for future testing, rather than testing ex ante hypotheses

Methods: Total effect of deforestation on malaria

9. Hypothetically, how much lower would malaria have been in a world with no deforestation?
    a. N=all the children surveyed for malaria
    b. For each surveyed child n=1:N we use the baseline predictive model (i.e. at observed levels of deforestation) to predict the probability the child tests positive for malaria

ii. The sum of these probabilities should be very near the number of children who actually tested positive for malaria

c. We then predict the probability that the child would have tested positive for malaria in a counterfactual world with 10% less/50% less/100% less deforestation
d. We divide the sum of predicted children testing positive for malaria in the counterfactual 10% less/50% less/100% less deforestation scenarios by the sum of predicted children testing positive for malaria in the baseline predicted scenario to obtain a statistic of how much lower malaria would have been with 10% less/50% less/100% less deforestation
e. Note: this assumes no difference in avoiding behavior. To the extent that less deforestation would lead to less malaria and less avoiding behavior, our estimates would be too small.

Methods: Marginal cost effectiveness analysis [D]

f. How cost-effective ($/DALY) is forest conservation as an anti-malarial intervention relative to other anti-malarial interventions?
    a. We want to calculate $/DALY averted from reducing deforestation, to compare with $/DALY averted from other anti-malarial interventions e.g. compared to White et al, Malaria Journal, 2011:
        i. $24/DALY for intermittent preventive treatment (IPT)
        ii. $27/DALY for insecticide-treated nets (ITN)
        iii. $143/DALY for indoor residual spraying (IRS)
        iv. Note: We need to inflate White's 2011 values to 2014$ values for comparison to Busch and Engelmann dollar values
        v. Note that these costs may not be static due to e.g. increasing drug resistance over time
    b. This will be crude, as it combines data from many disparate sources with much parameter uncertainty.
    c. We use a simulation, as follows (shown in Table 2 below).
        1. For every observation we "add" $100/ha/yr of forest conservation to the cell-year and estimate the resulting percentage change in deforestation (taken from Busch and Engelmann, 2015)
            a. Note that because Busch and Engelmann estimated how prices influence deforestation over 2001-2012 we must inflate prices slightly to produce the same changes over the 2001-2014 period
        2. We then calculate the additional probability that a surveyed child within that cell-year has malaria by calculating the difference in predicted malaria between the actual deforestation and the marginally lower deforestation, using the pooled malaria regression with covariates, based on the combined land cover change effect and land cover effect.
        3. We assume that the percentage change in malaria prevalence among children under 5 is equal to the

percentage change in lost DALYs from malaria among children under 5

4. To translate percentage changes into absolute changes, we:
    a. multiply by the total number of lost DALYs from malaria in children under 5 across the aggregate of all surveyed country-years
    b. And, divide by the total hectares of deforestation across the aggregate of all surveyed country-years

d. These operations result in (absolute reduction in lost DALYs from malaria in children under 5)/($ to achieve this reduction in deforestation)
e. Selected notes
    1. We could potentially apply a discount rate to convert the future flow of $/yr to a NPV of $. But, there's also a future flow of anti-malarial benefits from stopping deforestation that we'd want to consider too. This is complicated, so we just compare annual cost to annual benefit.
    2. Reducing deforestation has other benefits besides just reducing malaria incidence, whereas many medical interventions (treatment, bed nets) might only deal with malaria. If we were able to account for these non-malaria benefits, then the total social value of reducing deforestation relative to other interventions would be even higher.
    3. Surveys for malaria likely occurred in places with higher-than-average malaria areas, by design. Specifically, "[u]nlike the DHS, which is carried out at various times during the year, the MIS [Malaria Indicator Surveys] is usually timed to correspond with the high malaria transmission season. This is essential if the MIS includes biomarker testing for malaria." (http://dhsprogram.com/what-we-do/survey-types/mis.cfm) Thus if we were to extrapolate to other areas beyond our survey, the DALY/$ would likely be lower.
    4. DALY could potentially be multiplied by $/DALY to determine a cost-benefit ratio (or to obtain benefit($) per hectare). We don't find $/DALY numbers and we choose not to undertake this primary research given the multitude of complexities and ethical judgments.
    5. Targeted interventions *within countries* could be more cost-effective than the average by targeting where (malarial) population is high and forest cover is high
    6. High $/DALY might not be competitive with alternative interventions, but might still pass a benefit-cost ratio test

Table 1. Survey waves included and relevant sample sizes

Region	DHS country ID	DHS year	Fever	Rapid test	Microscopy
Africa	Angola	2006	736	0	0
Africa	Angola	2011	4 629	1 804	1 804
Africa	Burkina Faso	2003	7 502	0	0
Africa	Burkina Faso	2010	9 677	4 165	4 149
Africa	Burkina Faso	2014	5 230	4 551	4 531
Africa	Benin	2001	3 139	0	0
Africa	Benin	2012	7 543	2 094	2 071
Africa	Burundi	2010	5 811	0	0
Africa	Burundi	2012	3 299	2 601	2 905
Africa	Congo Democratic Republic	2007	4 381	0	0
Africa	Congo Democratic Republic	2013	10 388	4 620	4 600
Africa	Cote d'Ivoire	2012	4 149	1 784	1 721
Africa	Cameroon	2004	3 885	0	0
Africa	Cameroon	2011	5 752	0	0
Africa	Egypt	2003	3 200	0	0
Africa	Egypt	2005	7 852	0	0
Africa	Egypt	2008	6 312	0	0
Africa	Egypt	2014	8 516	0	0
Africa	Ethiopia	2005	7 439	0	0
Africa	Ethiopia	2010	8 517	0	0
Africa	Gabon	2012	1 937	0	0
Africa	Ghana	2003	2 444	0	0
Africa	Ghana	2008	1 780	0	0
Africa	Ghana	2014	3 117	1 360	1 360
Africa	Guinea	2005	4 042	0	0
Africa	Guinea	2012	4 303	1 977	1 977
Africa	Kenya	2003	3 690	0	0
Africa	Kenya	2008	4 157	0	0
Africa	Kenya	2014	12 928	0	0
Africa	Liberia	2007	3 129	0	0
Africa	Liberia	2009	2 078	1 849	1 851
Africa	Liberia	2011	1 689	1 482	1 415
Africa	Liberia	2013	4 536	0	0
Africa	Lesotho	2004	2 418	0	0
Africa	Lesotho	2009	2 713	0	0
Africa	Lesotho	2014	1 963	0	0

Africa	Morocco	2003	3 165	0	0
Africa	Mali	2001	8 261	0	0
Africa	Mali	2006	8 318	0	0
Africa	Mali	2012	6 929	3 204	3 181
Africa	Malawi	2004	8 493	0	0
Africa	Malawi	2010	15 882	0	0
Africa	Malawi	2012	1 566	1 368	1 362
Africa	Malawi	2014	1 380	1 197	1 210
Africa	Mozambique	2011	6 631	2 943	2 933
Africa	Nigeria	2003	3 096	0	0
Africa	Nigeria	2008	17 786	0	0
Africa	Nigeria	2010	3 677	3 133	3 124
Africa	Nigeria	2013	18 199	0	0
Africa	Namibia	2006	2 528	0	0
Africa	Namibia	2013	2 321	0	0
Africa	Rwanda	2005	5 972	0	0
Africa	Rwanda	2008	3 697	3 575	0
Africa	Rwanda	2010	6 946	3 128	3 165
Africa	Sierra Leone	2008	3 122	0	0
Africa	Sierra Leone	2013	7 028	0	0
Africa	Senegal	2005	5 892	0	0
Africa	Senegal	2008	10 078	2 616	2 688
Africa	Senegal	2010	7 344	2 195	2 184
Africa	Senegal	2012	4 133	3 572	3 550
Africa	Swaziland	2006	1 716	0	0
Africa	Togo	2013	4 480	2 068	2 076
Africa	Tanzania	2010	5 543	0	0
Africa	Uganda	2006	5 779	0	0
Africa	Uganda	2009	3 051	3 019	3 026
Africa	Uganda	2011	5 427	0	0
Africa	Zambia	2007	3 708	0	0
Africa	Zambia	2013	7 696	0	0
Africa	Zimbabwe	2005	3 402	0	0
Africa	Zimbabwe	2010	3 350	0	0
Americas	Dominican Republic	2007	4 499	0	0
Americas	Dominican Republic	2013	1 005	0	0
Americas	Guyana	2009	1 558	0	0
Americas	Honduras	2011	6 429	0	0
Americas	Haiti	2006	3 288	0	0
Americas	Haiti	2012	4 231	0	0
Americas	Peru	2004	4 698	0	0

Americas	Peru	2009	3 805	0	0
Asia	Bangladesh	2004	4 458	0	0
Asia	Bangladesh	2007	3 397	0	0
Asia	Bangladesh	2011	5 186	0	0
Asia	Bangladesh	2014	4 652	0	0
Asia	Indonesia	2003	8 341	0	0
Asia	Nepal	2001	5 453	0	0
Asia	Nepal	2006	3 935	0	0
Asia	Nepal	2011	3 776	0	0
Asia	Philippines	2003	3 596	0	0
Asia	Philippines	2008	3 532	0	0
Asia	Pakistan	2006	5 180	0	0
Asia	Timor-Leste	2009	7 043	0	0
	TOTAL		469 539	60 305	56 883

Table 2

	Coded as unfinished (0)	Coded as finished (1)
Floor	adobe bamboo clay dung earth mud palm sand wood planks	brick carpet cement ceramic concrete linoleum parquet polished wood tile vinyl
Wall	adobe bamboo bark cane earth grass mud palm sticks thatch wood wood planks	bricks cement concrete metal shingles stone
Roof	bamboo cana canvas cardboard carton estera grass mud mud bricks natte palm plastic tarp sod straw thatch tin cans wood wood planks	asbestos concrete finished wood iron metal sheets shingles tiles tin zinc
Water Source	Coded as collected from open: well borehole surface water spring river dam lake pond canal rainwater	Coded as piped or purchased: piped tap tank bottled water tanker truck peddler sachet bag vendor

	pump forage fountain	
Place of delivery	Coded as outside of health facility hospital health center health post clinic maternal and child welfare center medical college health complex doctor's office dispensary maternity home	Coded as in health facility respondent's home other's home traditional birth attendant midwife on the way to the hospital en route to provider

Table: $ - > reduced deforestation -> reduced malaria -> DALY averted

Operation	Value	Assumed value	Scale of value	Source
	$\frac{Average\ dollar\ value\ 2001-2012\ \left(2014\frac{\$}{ha}/yr\right)}{Average\ dollar\ value\ 2001-2014\ \left(2014\frac{\$}{ha}/yr\right)}$	1.021	Model-wide parameter	www.usinflationcalculator.com based on Consumer Price Index, US Department of Labor Bureau of Labor Statistics
Multiplied by…	$\frac{1\ percentage\ point\ reduction\ of\ deforestation\ (\%)}{\$\ to\ achieve\ 1\ percentage\ point\ reduction\ of\ deforestation\ 2001-2012\ \left(2014\frac{\$}{ha}/yr\right)}$	Decrease at $100/ha/yr Africa: 1.60%, Asia: 2.42%, Latin America: 0.98% Implies cost of reducing 1%: Africa: $62.31/yr, Asia: $41.03yr, LatAm: $102.05/yr	Continent-specific	Busch and Engelmann (2017).
Mulitiplied by…	$\frac{\%\ \boldsymbol{reduction\ in\ prob.a\ child\ has\ malaria\ when\ sampled}\ (\%)}{\boldsymbol{1\ percentage\ point\ reduction\ of\ deforestation}\ (\%)}$	**Predicted conditional value (land cover effect+land cover change effect)**	**Model aggregate**	**Model output**
Multiplied by…	$\frac{\%\ reduction\ in\ lost\ DALYs\ from\ malaria\ in\ children\ under\ 5\ (\%)}{\%\ reduction\ in\ prob.a\ child\ has\ malaria\ when\ sampled\ (\%)}$	Assumed to be equal to 1	Model-wide parameter	Assumption
Yields…	$\frac{\%\ reduction\ in\ lost\ DALYs\ from\ malaria\ in\ children\ under\ 5\ (\%)}{\frac{\$}{\frac{ha}{yr}}\ to\ achieve\ 1\ percentage\ point\ reduction\ of\ deforestation\ 2001-2014\left(2014\frac{\$}{ha}/yr\right)}$			
Multiplied by…	# lost DALYs from malaria in children under 5 within aggregate	Sum across surveys (country-years) Sensitivity: average across surveys	Aggregate [plus survey-wide as an exploratory analysis]	WHO
Divided by…	Hectares of deforestation in aggregate (ha)	Sum across surveys (country-years) Sensitivity: average across surveys	Aggregate [plus survey-wide as an exploratory analysis]	Hansen et al
Yields…	$\frac{absolute\ reduction\ in\ lost\ DALYs\ from\ malaria\ in\ children\ under\ 5\ (DALY)}{\$\ to\ achieve\ absolute\ reduction\ of\ deforestation\ \left(2014\frac{\$}{yr}\right)}$			