The effects of changes in distance to nearest health facility on under-5 mortality and health care utilization in rural Malawi, 1980-2000

Authors: John Quattrochi\textsuperscript{a*}, Kenneth Hill\textsuperscript{c}, Joshua A Salomon\textsuperscript{b}, Marcia C. Castro\textsuperscript{d}

Affiliations:
\textsuperscript{a}Simmons University, Department of Public Health, 300 The Fenway, Boston, MA, 02115, USA
john.quattrochi@simmons.edu
\textsuperscript{b}Harvard T.H. Chan School of Public Health, Department of Global Health and Population, 655 Huntington Ave., Boston, MA, 02115, USA
kenneth_hill_1@yahoo.com
\textsuperscript{c}Stanford University, Center for Health Policy and Center for Primary Care and Outcomes Research, 615 Crothers Way, Stanford, CA 94305, USA
salomon1@stanford.edu
\textsuperscript{d}Harvard T.H. Chan School of Public Health, Department of Global Health and Population, 655 Huntington Ave., Boston, MA, 02115, USA
mcastro@hsph.harvard.edu

*Corresponding author: john.quattrochi@simmons.edu
Abstract

Background: Despite important progress, the burden of under-5 mortality remains unacceptably high, with an estimated 5.3 million deaths in 2018. Lack of access to health care is a major risk factor for under-5 mortality, and distance to health care facilities has been shown to be associated with less access to care in multiple contexts, but few such studies have used a counterfactual approach to produce causal estimates.

Methods: We combined retrospective reports on 22,088 births between 1980 and 2000 from the 2000 Malawi Demographic and Health Survey with a 1998 health facility census that includes the date of construction for each facility, including 335 maternity or maternity/ dispensaries facilities built in rural areas between 1980 and 1998. Using a difference-in-differences approach with Cox proportional hazards models, we estimated the effect of a reduction in distance to nearest facility, conditional on initial distance, on under-5 mortality. Using a difference-in-difference approach with linear probability models, we estimated the effect of a reduction in distance to nearest facility on number of antenatal visits prior to delivery, place of delivery, receiving skilled assistance during delivery, and receiving a check-up following delivery.

Findings: We found no effects of a decrease in distance to the nearest health facility on the hazard of death before age five years. We also found no effect of reduced distance to nearest facility on utilization of maternal health services. The effects estimated here likely depend on the quality of health care, the availability of transportation, the demand for health services, and the underlying causes of mortality, among other factors.

Conclusion: Reducing under-5 mortality and increasing utilization of care in rural Malawi and similar settings may require more than the construction of new health infrastructure.

Keywords: under-5 mortality, utilization, maternal health, distance, service availability, Malawi
Background

Despite important progress, the burden of under-5 mortality remains unacceptably high, with an estimated 5.3 million deaths in 2018.\(^1\) The Sustainable Development Goals include ending preventable deaths of children under 5 by 2030.\(^2\) Lack of access to health care is a major risk factor for under-5 mortality, and distance to health care facilities has been shown to be associated with less access to care in multiple contexts.\(^3\)–\(^6\) Pregnant women living farthest from health facilities are generally the least likely to give birth in a facility, and least likely to utilize antenatal care.\(^3\),\(^5\)–\(^7\) Their children tend to be at higher mortality risk than those of mothers living closer to facilities. There remains a question, however, about the extent to which this relationship is causal.

Malawi has reduced its under-5 mortality rate from 250 per 1,000 live births in 1990 to 54 in 2018, a 78% reduction, despite an HIV/AIDS epidemic that peaked at approximately 15% prevalence.\(^8\) While health policy prior to Malawi’s first democratic elections in 1994 tended to favor private provision of health care, over the last 24 years Malawian policymakers have put forward bold programs to make health care more accessible, including 1994’s National Safe Motherhood Program and 2012’s Presidential Initiative for Safe Motherhood.\(^9\)–\(^15\)

In this paper, we took a difference-in-differences approach to estimate the causal effect of new facilities on under-5 mortality and utilization of maternal health services, namely, place of delivery, receiving a check-up following delivery, number of antenatal visits prior to delivery, or receiving skilled assistance during delivery.
Methods

Data

Health facility data. A national census of health facilities in 1998 carried out by the Malawi Ministry of Health and the Japanese International Cooperation Agency (MOH & JICA 1999) lists 719 facilities in operation. For most facilities, the data included the date of construction (589 out of the 719 (82%)), and for all facilities it included the GPS coordinates, the facility type, principle funder, and owner.

The dates of construction ranged from 1889 (when the first missionary hospital was built) to 1998. We restricted our analysis to the period from 1980 to 1998 due to limitations of the mortality data (described below). Of the 589 facilities for which we have date of construction, 337 (57%) were built between 1980 and 1998. There was no apparent temporal trend in the number of facilities built per year; the number ranged from a minimum of 10 in 1987 to a maximum of 32 in 1998 (Figure 1).

Of the 337 new facilities, 92% are one of two types: 135 (40%) are classified as “dispensary” and 175 (52%) are classified as “dispensary/maternity.” Dispensaries are permanent structures from which drugs are distributed. They provide outpatient care and may contain holding beds. Dispensary/maternities are similar but provide more extensive services to expectant mothers (antenatal, delivery, and postnatal care). The other facility types are district hospital, hospital, mental hospital, primary health center and urban health center. These are almost exclusively located in urban areas.
We assumed that those facilities missing the date of construction (130 out of 719), were built prior to 1980. If some of these facilities were in fact built between 1980 and 1998, then some births will be erroneously coded as being closer to a health facility than they actually were. This will bias the effect estimate towards the null if the true effect of being closer to a health facility is to reduce mortality or increase utilization.

*Mortality data.* The data on under-5 mortality came from the 2000 Malawi Demographic and Health Survey (MDHS), a nationally representative survey targeting all resident women aged 15-49. Variables collected include the date of birth and date of death (if applicable) for all children ever born to respondents (n=40,221 children). The data also included GPS coordinates for MDHS enumeration areas, which we refer to as villages in the remainder of the paper.

Migration has the potential to cause measurement error in the treatment variable since mothers’ residences in 2000 may not be in the same location as their residences in previous years. We restricted our analysis to rural births that occurred at the same location where the mother was living at the time of the survey, as reported in an MDHS question about length of time at one’s current residence.

**Operationalization of treatments and outcomes**

The primary outcome of interest was the hazard of death between birth and age five, estimated from retrospective birth histories in the MDHS that included the date of birth and date of death for each child. Children still alive and under age five at the time of the survey were right-censored (i.e. they have missing survival data between the age at which the survey occurred
The appendix includes tests showing that recall bias is not a concern in these data. We also estimated effects on the following secondary outcomes: (1) place of delivery, (2) receiving a check-up following delivery, (3) number of antenatal visits prior to delivery and (4) receiving skilled assistance during delivery.

For the mortality analysis, the treatment of interest is the reduction in distance to the nearest health facility caused by the construction of a new facility, conditional on distance to the nearest health facility prior to the construction of a new facility. We use multiple treatment variables to reflect the intuition that the benefit of a new facility depends on both the distance from the village to the old facility and the distance from the village to the new facility.

We created four distance categories (<2km, 2-5km, 5-10km, and >10km to nearest facility), which correspond to six possible treatments, each representing a move from one distance category to a nearer distance category: (1) >10km to 5-10km, (2) >10km to 2-5km, (3) >10km to <2km, (4) 5-10km to 2-5km, (5) 5-10km to <2km, and (6) 2-5km to <2km. The reference category is no change in distance category.

For each village we calculated the distance to the nearest health facility in each year from 1979 to 1998. We linked the change in distance category (including no change) for each village-year to each child-year in the mortality dataset. The dates of construction did not include the month of construction. Therefore, to avoid over-estimating exposure to new facilities, construction was assumed to have occurred on December 31. If a new facility was built during a child’s lifetime, we assigned the change in distance to the remaining person-time contributed by that child.
For the secondary outcomes, we operationalized the treatment variable differently than in the mortality analysis, due to much smaller sample sizes. The treatment was the reduction in distance to the nearest facility, *unconditional* on distance from the village to the nearest existing facility. We divided the treatment variable into three change-in-distance categories: less than or equal to 3km, 3-6km and more than 6km.

**Identification Strategy and Statistical Analysis**

An ideal study of the effects of new health facilities on under-5 mortality would randomly assign villages to receive a new health facility. By comparing the mortality rates before and after health facility construction in villages that did receive a facility to those that did not, the average treatment effect of a new facility could be easily calculated. In the current study, the location for new facilities may be endogenous to the under-5 mortality rate. If the placements were endogenous, and one were to simply carry out a cross-sectional comparison of areas with new health facilities to those without, it is unclear which direction the bias would take.

To overcome this potential endogeneity, we use a difference-in-differences design. Consider a simplified version of this study in which there are only two villages, A and B, two time periods, Pre and Post, and the effect of a new facility does not depend on the distance to the nearest existing facility. Village B receives a new health facility between the two time periods, while Village A does not. Let the hazard of death for child \( i \) in village \( v \) at time \( t \) be denoted \( \text{hazard}_{0|ivt} \), where the “0” indicates that it is a potential outcome if no new facility is built. Thus we will observe \( \text{hazard}_{0|ivt} \) for Village A but not for Village B. The key assumption to identify the causal effect of the new facility is that the hazard in Village B in the Post period *would have been the same* as the hazard in Village A in the Post period, had the new facility not been built.
Assuming also that the hazard can be modeled linearly, the regression to estimate the causal effect of the new facility is then:

\[
\text{hazard}_{ivt} = \alpha + \lambda B_v + \gamma \text{Post}_t + \delta (B_v \times \text{Post}_t) + \epsilon_{ivt}
\]

Where \(B_v\) is an indicator for Village B, \(\text{Post}_t\) is an indicator for the second time period, \((B_v \times \text{Post}_t)\) represents the treatment, and \(\epsilon_{ivt}\) is an error term. The effect of a new health facility, \(\delta\), is equal to the following difference in differences:

\[
\{E[\text{hazard}_{ivt} | v = B, t = \text{Post}] - E[\text{hazard}_{ivt} | v = B, t = \text{Pre}]\}
\]

\[
- \{E[\text{hazard}_{ivt} | v = A, t = \text{Post}] - E[\text{hazard}_{ivt} | v = A, t = \text{Pre}]\}
\]

The other coefficients are interpreted as follows:

\[
\alpha = \lambda A_v + \gamma \text{Pre}_t
\]
\[
\lambda = \lambda B_v - \lambda A_v
\]
\[
\gamma = \gamma \text{Post}_t - \gamma \text{Pre}_t
\]

With more than two villages and more than two time periods, the model becomes:

\[
\text{hazard}_{ivt} = \lambda_v + \gamma_t + \delta (\text{New facility}_v \times \text{Post}_t) + \epsilon_{ivt}
\]

where \(\lambda_v\) is a vector of village-specific dummies, \(\gamma_t\) is a vector of period-specific dummies, \((\text{New facility}_v \times \text{Post}_t)\) is the treatment, which indicates villages and periods after a new facility is built, and \(\epsilon_{ivt}\) is an error term. The key assumption of equivalence in expected outcomes in the absence of treatment remains. The hazard in the absence of treatment is assumed to be determined by the sum of a time-invariant village effect, \(\lambda_v\), and a period effect, \(\gamma_t\), that is common across villages:
This model is inappropriate for the current analysis due to three complications. First, we used multiple categories rather than one. As explained above, this reflects the intuition that the benefit of a new facility depends on both the distance from the village to the old facility and the distance from the village to the new facility. Second, the hazard was measured at the child level, rather than at the village level. Third, the child-level data contained right-censored observations (i.e. children who were under five and still alive at the time of the survey). For these reasons a survival model was more appropriate. We fit a semi-parametric Cox proportional hazards model, stratified by village:

\[ h(\text{age})_{ivy} = h(\text{age})_{0v} \exp[\alpha_y \text{Year of birth}_y + \delta_{1-6}(\text{New facility}_v \times \text{Post}_y) + \beta X_{ivt}], \]

where

\[ h(\text{age})_{0v} = \text{the baseline hazard of child death by age in village } v, \]

\[ \text{Year of birth}_y = 1 \text{ if child } i \text{ was born in year } y, \]

\[ \text{New facility}_v \times \text{Post}_y = 1 \text{ in any year } y \text{ that occurs after the construction of a new facility in village } v, \] which results in one of six possible changes in distance category, and

\[ X_{ivt} = \text{a vector of controls, namely indicators if the child is a first birth, mother’s age at birth, and mother’s education.} \]

The key assumption noted above about year and village effects now becomes one of multiplicity rather than additivity, and includes the unspecified baseline hazard:
\[ E[hazard_{0:t} | \nu, t] = H(\text{age})_{0:t} \exp(\gamma_t) \]

Another assumption, inherent in Cox models, is that the hazards are proportional. We tested this assumption in two ways. First, we tested for non-zero slope in a generalized linear regression of the scaled Schoenfeld residuals on time.\(^\text{20}\) Second, because that test can be “over-powered” – with many observations it may classify substantially insignificant changes in the hazard ratio as statistically significant -- we visually assessed plots of the scaled Schoenfeld residuals for the covariates that the test identified as violating the proportional hazards assumption.\(^\text{21}\)

P-values for distance reduction coefficients were adjusted for multiple testing using a false discovery rate procedure.\(^\text{22}\)

Ethical approval was obtained from Simmons University Institutional Review Board.

**Results**

There were a total of 40,306 births reported in the 2000 MDHS. Of those births, 22,088 were eligible for the mortality analysis, meaning that they occurred in a rural area between 1980 and 2000 to a woman in a village who reported living in the same village since at least the date of birth (Table 1). Of the mothers in the mortality analysis sample, 43.9% had no education, 54.9% had primary education, and 1.2% had secondary education or higher. For the utilization analysis, the number of eligible births ranged from 3,645 to 6,815.

**Table 1. Descriptive statistics**

| Primary outcomes, exposure, and control variables among eligible rural births |
|-------------------------------|-------------------|-----------------|
| Child deaths | Obs | Number of deaths | Percent |
|---------------|-----|-------------------|---------|
|                | 22088 | 4,297 | 19.5 |

| Obs | Mean | Std. Dev. | Min | Max |
|-----|------|-----------|-----|-----|
Distance to nearest health facility (meters) | 22088 | 6005 | 4559 | 305 | 47442
--- | --- | --- | --- | --- | ---
Births to women with no education | 22088 | 9241 | 41.8 |
Births to women with primary education | 22088 | 12469 | 56.5 |
Births to women with secondary education or higher | 22088 | 378 | 1.7 |
First birth | 22088 | 4467 | 20.2 |
Twin | 22088 | 871 | 3.9 |
Births to a mother under 19 years old | 22088 | 4247 | 19.2 |
Births to a mother over 35 years old | 22088 | 2633 | 11.9 

*Secondary outcomes (only asked to mothers of children under five years old)*

| Outcome | Obs | Number | Percent |
| --- | --- | --- | --- |
| Home birth | 6768 | 3343 | 49.4 |
| At least three antenatal visits prior to delivery | 4777 | 2589 | 54.2 |
| Skilled assistance during delivery | 6815 | 3380 | 49.6 |
| Check-up following delivery | 3645 | 259 | 7.1 |

Notes: data from 2000 Malawi Demographic and Health Survey. Eligible rural births are those that occurred between 1980 and 2000 to a woman who reports having lived in the same village when she gave birth as she did at the time of the interview.

Kernel density plots of births by distance to nearest facility for selected years show that the average distance has been decreasing over time, driven in particular by facilities 10-20km from a village being replaced by facilities less than 10km away from a village (Figure 3). Overall, the modal distance is 5-10km, and very few births occurred within 1km of a health facility (Figure 4). Turning to the number of births per year, the number is higher in more recent years because the population is growing rapidly and the data come from women who were aged 15-45 in 2000. The sharp decrease in births in 1995 is almost certainly due to misreporting of four year olds as five year olds. This means that for a small number of child-years, children will be coded as being untreated when they were in fact treated. If the true effect of health facilities is protective, then this will bias estimates toward the null.

[FIGURE 3 HERE]

[FIGURE 4 HERE]

Each village contributed an average of 1,993 child-months (166 child-years) to the analysis (min 90; max 5,341) (Table 2). On average, 1,483 of those child-months (74%) occurred
prior to the construction of a new facility. Of the child-months that occurred after the
construction of a new facility, relatively few were contributed by villages in which the distance
to nearest facility decreased to less than 2km (an average of 88 child-months per village, or 4.4%
of all child-months in the analysis). Thus the statistical analysis that follows may have relatively
little power to detect effects from changes to less than 2km.
We present results from four stratified Cox proportional hazards models (Table 3; additional specifications are presented in the appendix). All four models stratified the baseline hazard by village (n=449) and include dummy variables for year of birth (n=19). Model 1 includes the six original treatment variables without covariates. None of the decreases in distance produce a statistically significant reduction in the hazard of under-five mortality at the 0.05 significance level, even before adjustment for multiple hypothesis testing.

Model 2 adds control variables and retains the six treatment variables. Again, none of the treatment variables produced significant effects. Model 3 includes a new coding of the treatment variable that collapsed reductions in distance from over 2km to less than 2km into one category. We found no effect from this reduction in distance. Model 4 included another coding of the treatment variable that collapsed all categories into one, namely more than 5km to less than 5km. Again, we found no effect on the hazard of death. The appendix includes our tests of the proportional hazards assumption, which we concluded was reasonable for this analysis.
Table 3. The effect of increased access to health care on under-5 mortality in rural Malawi, 1980-2000

| Treatment or control variable | Model 1          |          | Model 2          |          | Model 3          |          | Model 4          |          |
|-------------------------------|------------------|---------|------------------|---------|------------------|---------|------------------|---------|
|                               | Hazard ratio     | adj. p- | Hazard ratio     | adj. p- | Hazard ratio     | adj. p- | Hazard ratio     | adj. p- |
|                               | (95% CI)         | value   | (95% CI)         | value   | (95% CI)         | value   | (95% CI)         | value   |
| >10km to 5-10km               | 1.013            | 0.96    | 1.025            | 0.96    | 1.025            | 0.96    | 1.025            | 0.96    |
|                               | (0.801,1.282)    |         | (0.810,1.298)    |         | (0.810,1.298)    |         | (0.810,1.298)    |         |
| >10km to 2-5km                | 1.233            | 0.73    | 1.256            | 0.73    | 1.256            | 0.73    | 1.256            | 0.73    |
|                               | (0.859,1.770)    |         | (0.871,1.810)    |         | (0.871,1.810)    |         | (0.871,1.810)    |         |
| >10km to <2km                 | 1.101            | 0.96    | 1.046            | 0.96    | 1.046            | 0.96    | 1.046            | 0.96    |
|                               | (0.610,1.989)    |         | (0.580,1.885)    |         | (0.580,1.885)    |         | (0.580,1.885)    |         |
| 5-10km to 2-5km               | 0.806            | 0.52    | 0.79             | 0.52    | 0.79             | 0.52    | 0.79             | 0.52    |
|                               | (0.637,1.021)    |         | (0.623,1.003)    |         | (0.623,1.003)    |         | (0.623,1.003)    |         |
| 5-10km to <2km                | 0.898            | 0.87    | 0.880            | 0.85    | 0.880            | 0.85    | 0.880            | 0.85    |
|                               | (0.625,1.289)    |         | (0.613,1.263)    |         | (0.613,1.263)    |         | (0.613,1.263)    |         |
| 2-5km to <2km                 | 1.406            | 0.73    | 1.453            | 0.73    | 1.453            | 0.73    | 1.453            | 0.73    |
|                               | (0.724,2.731)    |         | (0.751,2.814)    |         | (0.751,2.814)    |         | (0.751,2.814)    |         |
| >2km to <2km                  |                  |         | 1.006            | 0.96    |                  |         | 0.933            | 0.83    |
|                               |                  |         | (0.761,1.330)    |         |                  |         | (0.790,1.102)    |         |
| >5km to <5km                  | 1.049            | 0.312   | 1.048            | 0.316   | 1.048            | 0.316   | 1.048            | 0.316   |
|                               | (0.957,1.149)    |         | (0.956,1.149)    |         | (0.956,1.149)    |         | (0.956,1.149)    |         |
| Twin                          | 3.008            | <0.001  | 3.001            | <0.001  | 3.003            | <0.001  | 3.003            | <0.001  |
|                               | (2.677,3.380)    |         | (2.671,3.372)    |         | (2.672,3.374)    |         | (2.672,3.374)    |         |
| Mother under 19               | 1.363            | <0.001  | 1.362            | <0.001  | 1.362            | <0.001  | 1.362            | <0.001  |
|                               | (1.245,1.493)    |         | (1.244,1.492)    |         | (1.244,1.492)    |         | (1.244,1.492)    |         |
| Mother over 35                | 1.010            | 0.847   | 1.010            | 0.852   | 1.009            | 0.859   | 1.009            | 0.859   |
|                               | (0.911,1.120)    |         | (0.911,1.120)    |         | (0.910,1.119)    |         | (0.910,1.119)    |         |
| Mother has primary education  | 0.882            | <0.001  | 0.882            | <0.001  | 0.881            | <0.001  | 0.881            | <0.001  |
|                               | (0.822,0.946)    |         | (0.822,0.946)    |         | (0.822,0.945)    |         | (0.822,0.945)    |         |
| Mother has secondary education| 0.474            | <0.001  | 0.475            | <0.001  | 0.475            | <0.001  | 0.475            | <0.001  |
|                               | (0.324,0.692)    |         | (0.325,0.694)    |         | (0.325,0.693)    |         | (0.325,0.693)    |         |
| Observations                  | 22,088           |         | 22,088           |         | 22,088           |         | 22,088           |         |

Notes: Hazard ratios from Cox proportional hazards models, with baseline hazard stratified by village of birth (n=449) and with fixed effects for year of birth. P-values for distance reduction coefficients were adjusted for multiple testing using the Benjamini and Hochberg (1995) false discovery rate procedure.

We found no effects of reductions in distance to the nearest facility on place of delivery, post-delivery checkup, three or more antenatal visits, or skilled assistance at delivery (Table 4).
Table 4. Effects of reductions in distance to nearest facility on maternal health care utilization

| Change in distance to nearest facility | Home birth | Post-delivery checkup | Antenatal visits | Skilled assistance at delivery |
|---------------------------------------|------------|-----------------------|------------------|------------------------------|
| <3km reduction                        | 0.05 (-0.04, 0.15) | 0.01 (-0.07, 0.10) | -0.05 (-0.15, 0.05) | -0.01 (-0.16, 0.13) |
| 3-6km reduction                       | -0.01 (-0.14, 0.12) | -0.10 (-0.20, 0.01) | 0.03 (-0.10, 0.17) | -0.11 (-0.34, 0.12) |
| >6km reduction                        | 0.03 (-0.14, 0.20) | -0.11 (-0.24, 0.02) | -0.03 (-0.20, 0.14) | -0.05 (-0.28, 0.19) |

Notes: Coefficients from linear probability models with fixed effects for village and month-year of birth (n=63), and the following control variables: child is first birth, child is twin, mother <19 yrs old, mother >35 yrs old, mother has primary education, mother has secondary education. None of the effects are significantly different from zero at the 0.05 level.

Discussion

Key results

This study found no evidence for the claim that the construction of new health facilities in rural Malawi between 1980 and 2000 caused a reduction in under-5 mortality. Nor did we find any evidence that number of antenatal check-ups, skilled assistance at birth, delivery in a facility, or post-natal check-ups increased following the construction of a new facility.

Limitations

There are several important limitations to this study. First, new health facilities were not constructed in randomly assigned locations. Therefore, it is possible that children born in areas where new facilities were built were systematically more or less likely to die before age five than children in other areas. We used a difference-in-differences estimation strategy to overcome this endogeneity, but it relied on the assumption that the underlying village-level hazard of under-5 mortality can be modeled as a multiplicative combination of time-invariant village effects and year-specific effects that are common across villages. That assumption implies that, after adjusting for mother’s education, age at birth of her child, whether or not the child is first born,
and village characteristics that do not change during the study period, secular trends in under-5 mortality are equivalent across villages.

Second, while we estimated the average treatment effect, new health facilities are likely to produce heterogeneous effects depending on the quality of care provided. The training of the facility staff, the frequency with which they work, and the availability of pharmaceuticals and other medical supplies may vary widely across facilities. We restricted the analysis to dispensaries and maternity/dispensaries, but even within these categories the variability in quality of care may be significant. Variability in quality of care over time – if, for example, new and better health care technologies become available in more recent years – can also lead to heterogeneous effects for similar reasons.

Third, we had limited data on utilization. The MDHS only included those variables for births in the five years prior to the survey, so the sample size was roughly one-quarter of that for the mortality analysis. It remains possible that utilization of those services increased in earlier years, or that utilization of other services (e.g., treatment for diarrhea, pneumonia or malaria) increased.

Fourth, the mortality data were retrospective accounts of women who were 15-49 years old in 2000. Children born to women who died were not represented. Because a child is more likely to die if her mother dies, retrospective estimates of under-5 mortality will tend to be underestimates if a substantial proportion of women die between the ages 15 and 49. The potential for bias in this study depends on two factors: the impact of new facilities on adult mortality, and the propensity for new facilities to be built in areas with increasing or decreasing adult mortality. Assume first that new facilities have no impact on adult or child mortality. If
new facilities were built in areas where adult mortality is increasing and, thus, under-5 mortality is underestimated, then it will appear that the new facility caused a reduction in under-5 mortality. If, on the other hand, new facilities were built in areas where adult mortality was decreasing, then the opposite will occur. The onset of the HIV epidemic in Malawi in the mid-1980s increased adult mortality substantially.\textsuperscript{23} It is unlikely, however, that new facilities were targeted to areas with higher or lower HIV prevalence. As we argued above, the information and implementation capacity necessary to target facilities based on burden of disease were not available. The availability of antiretroviral therapy (ART) also impacts adult mortality. In Malawi, ART was not widely available in until after 2004,\textsuperscript{24} thus it would not affect the analysis presented here.

**Interpretation**

Previous studies on the relationship between distance to nearest health facility and child mortality show mixed results. A study that combined survey data from 21 countries found that greater distance was associated with higher neonatal mortality, but not mortality at higher ages.\textsuperscript{25} A study in a rural Kenyan districts with high health facility density found no association between travel time and child mortality,\textsuperscript{26} nor did a case-control study in rural Gambia.\textsuperscript{27} The file drawer problem suggests that studies finding no relationship between distance and mortality are less likely to be published.\textsuperscript{28}

Studies reporting an association between distance and mortality include a study in rural Burkina Faso which estimated that under-5 mortality was 50% higher at a distance of 4 hours walking time to the nearest facility compared with having a facility in the village.\textsuperscript{4} A matched pairs study found that the construction of maternity clinics in Indonesia in the mid-1980s reduced infant mortality in the surrounding area by 15%.\textsuperscript{29} In South Africa, allowing blacks to utilize
facilities that were formerly restricted to whites increased the weight-for-age scores for male infants, but had no effect on female infants.\textsuperscript{30} Several other observational studies have found a positive association between distance and under-5 mortality.\textsuperscript{31–33}

**Generalizability**

The external validity of this study benefits from the fact that it used data from villages and health facilities throughout rural Malawi, over a period of twenty years. Nevertheless, this is a study of a single country during a particular phase of history. As discussed above, the effects estimated here likely depend on the quality of health care, the availability of transportation, the demand for health services, and the underlying causes of mortality, among other factors. Data on these factors are scarce in Malawi and other high mortality countries, and more resources should be invested in collecting that information. The results presented here suggest that there is more to the story of reducing under-5 mortality than increasing the availability of health infrastructure by shortening travel distances. This finding may hold across other low-income, high-mortality countries, particularly in sub-Saharan Africa. More research is needed on the relationship between access to care, quality of care, and perceived quality of care in these settings.

**List of abbreviations**

ART = antiretroviral therapy

MDHS = Malawi Demographic and Health Survey
Declarations

Ethics approval and consent to participate

Ethical approval was obtained from Simmons University Institutional Review Board.

Consent for publication

Not applicable.

Availability of data and materials

Data and code are available from the corresponding author upon reasonable request.

Competing interests

The authors declare they have no competing interests.

Funding

JQ was supported by a Harvard University Presidential Scholarship and an NIH Infectious Disease and Biodefense traineeship (T32 AI007535, PI: George Seage). The funders had no role in the design of the study; collection, analysis, and interpretation of data; or writing the manuscript.

Authors’ contributions

JQ’s role included conceptualization, data curation, formal analysis, software, and writing the original draft. JS, KH, and MC’s roles included conceptualization and editing and review of the manuscript. All authors read and approved the final manuscript.

Acknowledgements
We thank Patrick Gerland for sharing the health facility data. We acknowledge the contributions of Matthew Kraft, Richard Murnane, and John Willett to this paper. All mistakes are our own.
References

1 UNICEF Child Mortality Data. UNICEF DATA. https://data.unicef.org/topic/child-survival/under-five-mortality/ (accessed Feb 18, 2020).

2 Sustainable Development Goals. https://sustainabledevelopment.un.org/sdgs (accessed Feb 18, 2020).

3 Lohela TJ, Campbell OMR, Gabrysch S. Distance to care, facility delivery and early neonatal mortality in Malawi and Zambia. *PloS One* 2012; 7: e52110.

4 Schoeps A, Gabrysch S, Niamba L, Sié A, Becher H. The Effect of Distance to Health-Care Facilities on Childhood Mortality in Rural Burkina Faso. *Am J Epidemiol* 2011; 173: 492–8.

5 Gabrysch S, Cousens S, Cox J, Campbell OM. The influence of distance and level of care on delivery place in rural Zambia: a study of linked national data in a geographic information system. *PLoS Med* 2011; 8.

6 Nesbitt RC, Lohela TJ, Soremekun S, et al. The influence of distance and quality of care on place of delivery in rural Ghana. *Sci Rep* 2016; 6: 1–8.

7 Hanson C, Gabrysch S, Mbaruku G, et al. Access to maternal health services: geographical inequalities, United Republic of Tanzania. *Bull World Health Organ* 2017; 95: 810–20.

8 World Bank. World Bank Data Catalog. 2020 https://data.worldbank.org/indicator (accessed April 18, 2019).

9 Banda EEN, Simukonda HP. The public/private mix in the health care system in Malawi. *Health Policy Plan* 1994; 9: 63–71.

10 Cortez R, Saadat S, Chowdhury S, Sarker I. Maternal and Child Survival: Findings from five countries experience in addressing maternal and child health challenges. The World Bank, 2014 https://ideas.repec.org/p/wbk/hnpdps/91294.html (accessed Feb 18, 2020).

11 Ministry of Health and Population, Ministry of Women and Children Affairs. Making motherhood safe for Malawian women. Lilongwe, 1994.

12 Kulmala T, Vaahtera M, Rannikko J, et al. The relationship between antenatal risk characteristics, place of delivery and adverse delivery outcome in rural Malawi. *Acta Obstet Gynecol Scand Orig Artic* 2000; 79: 984–990.

13 Walsh A, Matthews A, Manda-Taylor L, et al. The role of the traditional leader in implementing maternal, newborn and child health policy in Malawi. *Health Policy Plan* 2018; 33: 879–87.

14 Doherty T, Zembe W, Ngandu N, et al. Assessment of Malawi’s success in child mortality reduction through the lens of the Catalytic Initiative Integrated Health Systems Strengthening programme: Retrospective evaluation. *J Glob Health* 2015; 5.
15Jahn A, Floyd S, Crampin AC, et al. Declining child mortality in northern Malawi despite high rates of infection with HIV. Bull World Health Organ 2010; 88: 746–753.

16National Statistical Office, ORC Macro. Malawi Demographic and Health Survey 2000. Zomba, Malawi and Calverton, Maryland, USA, 2001.

17Angrist J, Pischke J-S. Mostly Harmless Econometrics. Princeton University Press, 2009.

18Therneau TM. Modeling Survival Data: Extending the Cox Model. Springer, 2000.

19Cox DR. Regression Models and Life-Tables. J R Stat Soc Ser B Methodol 1972; 34: 187–220.

20Grambsch PM, Therneau TM. Proportional Hazards Tests and Diagnostics Based on Weighted Residuals. Biometrika 1994; 81: 515–26.

21Box-Steffensmeier JM, Jones BS. Event History Modeling: A Guide for Social Scientists. Cambridge University Press, 2004.

22Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Methodol 1995; 57: 289–300.

23UNAIDS. Report on the Global AIDS Epidemic. 2010.

24WHO, UNICEF, UNAIDS. Global update on HIV treatment 2013: Results, impact and opportunities. Geneva, 2013
http://www.who.int/hiv/pub/progressreports/update2013/en/index.html (accessed Sept 30, 2013).

25Karra M, Fink G, Canning D. Facility distance and child mortality: a multi-country study of health facility access, service utilization, and child health outcomes. Int J Epidemiol 2017; 46: 817–26.

26Moïsi JC, Gatakaa H, Noor AM, et al. Geographic access to care is not a determinant of child mortality in a rural Kenyan setting with high health facility density. BMC Public Health 2010; 10: 142.

27Rutherford ME, Dockerty JD, Jasseh M, et al. Access to health care and mortality of children under 5 years of age in the Gambia: a case-control study. Bull World Health Organ 2009; 87: 216–24.

28Rosenthal R. The file drawer problem and tolerance for null results. Psychol Bull 1979; 86: 638.

29Frankenberg E. The effects of access to health care on infant mortality in Indonesia. Health Transit Rev 1995; 15: 143–63.
30 Tanaka S. Does abolishing user fees lead to improved health status? Evidence from post-Apartheid South Africa. *Am Econ J Econ Policy* 2014; 6: 282–312.

31 Schellenberg JRA, Mrisho M, Manzi F, et al. Health and survival of young children in southern Tanzania. *BMC Public Health* 2008; 8: 194.

32 Van den Broeck J, Eeckels R, Massa G. Maternal determinants of child survival in a rural African community. *Int J Epidemiol* 1996; 25: 998–1004.

33 Magnani RJ, Rice JC, Mock NB, Abdoh AA, Mercer DM, Tankari K. The impact of primary health care services on under-five mortality in rural Niger. *Int J Epidemiol* 1996; 25: 568–77.

**Figure legends**

**Figure 1. Number of new health facilities constructed per year in Malawi, 1980 to 1998**
Notes: data from national census of health facilities in 1998 carried out by the Malawi Ministry of Health and the Japanese International Cooperation Agency

**Figure 2. Map of health facilities and enumeration areas in Malawi**
Notes: map created by the authors using data from 1998 Malawi Health Facility Census and the 2000 Malawi Demographic and Health Survey

**Figure 3. Distribution of births by distance from village to nearest health facility, selected years of birth**
Notes: Kernel densities of birth in given year by distance from village to nearest health facility. Data from 2000 Malawi Demographic and Health Survey and 1998 Malawi Health Facility Census.

**Figure 4. Births per year by distance to nearest health facility in analysis sample**
Notes: Data from 2000 Malawi Demographic and Health Survey (DHS) and 1998 Malawi Health Facility Census. Age-heaping in 1994 likely due to the additional DHS questions required of children under 5 years old.