

A conversation with Dr. Kevin Croke on September 25th, 2014

Participants

- Dr. Kevin Croke — post-doctoral research fellow, Department of Global Health and Population, Harvard School of Public Health
- Jake Marcus — Research Analyst, GiveWell
- Milan Griffes — Research Analyst, GiveWell

Note: These notes were compiled by GiveWell and give an overview of the major points made by Dr. Croke.

Summary

GiveWell spoke with Dr. Croke about his research on the educational impacts of a randomized controlled trial (RCT) of a deworming program in Uganda. Conversation topics included the robustness of the results, plausibility of the treatment-on-the-treated estimate, the external validity of his results, potential publication bias in research on deworming RCTs, applicability of his research methodology to other deworming RCTs, and initial responses to his paper.

Background information

The Tanzanian non-governmental organization Twaweza's research initiative, Uwezo, surveys literacy and numeracy of children in Kenya, Tanzania, and Uganda. After speaking with the World Bank economist Owen Ozier, Dr. Croke decided to use Uwezo literacy and numeracy data to follow up on an RCT of a deworming program for school children in Uganda. Uwezo data was collected seven to eight years after the original RCT, and was available for about half of the treatment and control parishes (a local administrative division in Uganda). Dr. Croke found significantly better scores in literacy and numeracy among children in the treated parishes compared to the control parishes, when controlling for age, gender, survey year, and interactions of these variables.

Robustness checks

When testing for statistical significance using regression equations, it is common practice to vary the control and dependent variables to other plausible specifications. If the effect cannot be observed under other plausible regression specifications, the result is not robust.

Overall, the observed effects of the deworming RCT in Uganda were robust to other plausible specifications of dependent and control variables:

- **Dependent variables** – The results are statistically significant using either raw scores or standardized test scores as the dependent variable. A binary variable, like passing a time-telling test, could have been used as the dependent variable, but would be less informative than a continuous variable. (Also it is not clear that telling time is a reflection of cognitive ability, compared to reading or doing math.)

- **Control variables** – When control variables for age, gender, survey year, and interactions between these variables were included in the regression equation, the observed differences in educational outcomes for treatment and control parishes were statistically significant. Without controlling for any variables, the differences were not statistically significant. However, it is standard practice to control for age since age is highly correlated with educational outcomes (i.e. a twelve year old is likely to be better at reading and math than a seven year old). Survey year was included as a control variable, but it is not likely to have a large impact since most observations are from 2011. Dr. Croke tested many different specifications for the regressions (not all of them reported in the working paper) and adding control variables only strengthened the conclusions. However, alternative specifications for dependent and control variables are fairly limited, since Uwezo surveys only include a few questions about socioeconomic status and other variables.

The marginal statistical significance for the unadjusted results is probably on account of the low sample size of the study. It is common practice to adjust for variables orthogonal to the treatment to soak up some of the variation and give greater power to detect an effect.

Potential confounding variables

Neither of these potential confounding variables appear to have biased the results:

- **Mobilization campaigns** – Mobilization campaigns (efforts aimed at increasing school attendance) were often conducted for Child Health Days. Since Child Health Days were a part of the original deworming project and therefore were implemented in both treatment and control parishes, there is no reason to believe that mobilization campaigns occurred more often in treatment or control parishes.
- **Migration** – It is unlikely that migration biased the results of this study. The children surveyed were still too young to have moved away from home. If some migration occurred, it would probably be among people older than the children studied here and would mostly be to Kampala or maybe across the border to Kenya. It is unlikely that migration was unbalanced between treatment and control parishes, since the parishes were randomly assigned during the original RCT.

Plausibility of the treatment-on-the-treated estimate

The regression equation measures an intent-to-treat (ITT) effect, since it estimates the effect of living in a treatment parish, not the effect of actually receiving deworming medication. Due to absence from school on a Child Health Day, some children in treatment parishes may not have actually received deworming medication. Dr. Harold Alderman, who was a principle investigator on the original deworming RCT, estimates that between 2/3^{rds} and 3/4^{ths} of children in treatment parishes actually received deworming medication in a given year.

The effect of actually receiving the treatment would be larger than the ITT estimate, since the average test scores in treatment parishes would be dragged down by those who did not

receive the treatment. Dr. Croke's adjustments for program coverage in treatment parishes and an increase in voluntary deworming medication use in control parishes estimate the treatment-on-the-treated (TOT) effect, which was close to a full standard deviation.

Dr. Croke believes that the TOT effect is surprisingly large. A speculative explanation of the effect size comes from the epidemiology of the disease. In the treatment parishes, all children in school were treated at the same time, which could completely wipe out worms in the area until they were brought in again by an outside visitor. In control parishes, some families purchased deworming drugs for their children independently, but there was still a consistent reservoir of worms in the community. The additional benefits of herd immunity achieved when everyone is treated for worms at the same time may explain why the TOT effect is so large.

External validity of the results

External validity refers to the extent to which results from a study can be generalized to other populations. If some conditions within a given study are unusual, other populations might not have the same results from the intervention. Dr. Croke's study may have some of the same external validity issues as Miguel and Kremer 2004, a study on a deworming RCT in Kenya. Due to an El Niño year, flooding caused unusually high infection rates during the years that some schools were treated with deworming medication in Kenya. Since some of the parishes in the Uganda RCT were close to Lake Victoria and the schools in Kenya, it is possible that they were affected by El Niño and had unusually high infection rates as well. Dr. Harold Alderman estimated that infection rates in Uganda were higher than average as reported in a Cochrane meta-analysis, but not as high as those in Kenya.

Uganda and Kenya both have high rates of malaria. Deworming medication could hypothetically have some effect on malaria incidence, which could also be affecting the results observed in both countries. However, a connection between deworming medication and malaria incidence is purely speculative.

Issues with publication bias, and applicability of the methodology to other contexts

It is possible that another researcher also used a later educational survey as a follow-up to a deworming medication RCT, but did not publish the results because no effect was found. However, Dr. Croke is not aware of anyone else who has done this kind of research, or of any other papers on deworming that GiveWell is not already aware of.

Dr. Croke considered using Uwezo data as another follow up for the deworming RCT in Kenya. However, the RCT in Uganda was randomized at the parish level, whereas the RCT in Kenya was randomized at the school level. Since the Uwezo data is organized at the local administrative level it is difficult to match the RCT treatment and control schools with the Uwezo data in Kenya. Attempting to match the Uwezo data to the school level RCT data would likely result in false positives and false negatives, which could bias the results.

It might be possible to employ Dr. Croke's methodology to RCTs of deworming in India and later educational data sources, like the Deworming and Enhanced Vitamin A study or the National Family Health Survey.

Initial responses to the paper

Dr. Croke has published the paper on his website, presented it at Harvard, and sent it to a few colleagues who are interested in deworming, including Dr. Harold Alderman. Overall, the responses to the paper were positive. Some colleagues suggested a few extra robustness checks, and methods for presenting the paper to a public health audience. The paper has not yet been submitted for peer review.

All GiveWell conversations are available at <http://www.givewell.org/conversations>