

This document describes the rationale for the decay adjustment in our deworming cost-effectiveness analysis. We have incorporated this adjustment thanks to [criticism](#) from the Happier Lives Institute.

In a nutshell

- The main piece of evidence we use for the long-term effects of deworming is an RCT in Kenya with follow-ups at ~10 years (KLPS-2), ~15 years (KLPS-3) and ~20 years (KLPS-4) after children received deworming treatment. While these surveys show decline in effect on ln earnings and consumption over time, we have typically viewed the different estimates across surveys as noisy estimates of the same effect and assumed the effects of deworming are constant throughout a person's working life.
- We now think we should account for some decay in benefits over time. We incorporate this decay by making the following key assumptions:
 - We put 50% weight on the interpretation that the different estimates over time are capturing true differences in effect size. While the data point to an estimate of decline, the confidence intervals are wide and there are differences in how data were collected over time, which make us reluctant to put full weight on KLPS 2-4 capturing true decay over time.
 - We set a prior that the effects are constant over time. This is based on a shallow literature review of studies of interventions during childhood where researchers reported at least two follow-ups on income during adulthood. We find a similar number of studies finding a decline in effect as an increase in effect over time.
 - We update from that prior at each time period (10-years, 15-years, and 20-years), using the informal Bayesian adjustment approach we've used previously.
 - We then extrapolate effects through the rest of the individual's working life based on the measured decline from 10-year to 20-year follow-up.
- Our best guess is that we should apply a -10% adjustment due to the possibility of decay in effects over time. While the decline in effects in later years leads to lower cost-effectiveness, this is partially counterbalanced by higher estimated effects in earlier years and by our putting only 50% weight on the interpretation that declines in measured effects across follow-ups reflects a true decline in effect over time.
- We have several uncertainties about this analysis:
 - This decay adjustment builds on top of our current Bayesian approach for estimating the effect of deworming. As a result, it's subject to the same limitations of that approach. It's possible that in the future we should overhaul our approach, which could lead to meaningful differences in how we incorporate decay.
 - The model is sensitive to our prior on whether effects should decay or not, and our current prior is based on a shallow literature review. If we expected effects to decay, we would include a stricter adjustment because we would (i) be updating from a prior where decay was already occurring and (ii) put more weight on the decay interpretation. We could potentially refine this estimate with a more thorough review of the literature and additional data analysis.

- The weight we put on whether these are noisy estimates of the same effect or different effects over time is based on a qualitative and highly subjective assessment. Putting higher weight on the surveys capturing different effects over time, for example, would lead to a stronger discount.

What we did previously

The main piece of evidence we use for the long-term effects of deworming is an [RCT](#) in Kenya that measures effects on income at ~10 years (KLPS-2), ~15 years (KLPS-3) and ~20 years (KLPS-4) after children receive deworming treatment.¹

Our typical approach has been to pool effects on earnings and consumption across three survey rounds, which suggests an effect of [0.109](#) on ln income.

Because deworming has limited high-quality evidence for an impact on income, we substantially discount this observed effect from the three survey rounds.² Our prior is that a plausible effect of deworming in the RCT in Kenya is [~1%](#).³ The RCT evidence, which finds an effect of [~10%](#), updates us slightly from that prior.⁴ Using an informal Bayesian updating framework, our best guess is that the effect for individuals in the RCT is [~1.4%](#), i.e., we apply a replicability adjustment of [13%](#) to the findings from the RCT in Kenya.⁵

We then assume that any effects of deworming last for [40 years](#) once an individual enters the labor force (assumed to be [8 years](#) after receiving deworming treatment). We assume these follow-ups provide noisy estimates of the same effect, and our prior is that effects should be constant over this 40-year period.

Incorporating the possibility that there is decay over time

An alternative interpretation is that the estimates across surveys reflect different effects of childhood deworming over time. If we take the survey estimates at face value, there appears to be a [decline in effect over time](#) (0.234 to 0.069 to 0.039 in ln earnings, KLPS-2 to KLPS-4, and 0.30 to 0.09 in ln consumption, KLPS-3 to KLPS-4).

¹ “Wage earnings and self-employment profits were collected in KLPS-2, KLPS-3, and KLPS-4; agricultural profits were collected in KLPS-3 and KLPS-4. Annual per capita household earnings are calculated as the sum of wage employment earnings, self-employment profits, and agricultural profits across all household members, divided by the number of household members. Household earnings are only available in KLPS-4.” [Hamory et al. 2021](#), Table 1.

² We describe the rationale for this [here](#) and [here](#).

³ This is based on evidence from health and other possible mechanisms that might contribute to deworming’s long term effects. Our calculations are in [this spreadsheet](#). 1% is the weighted average of effects from different mechanisms ([these cells](#)) with the weights on these different mechanisms ([these cells](#)).

⁴ The treatment effect of deworming on ln(income) in the Miguel and Kremer 2004 study population is 0.109, based on our pooling of results across rounds. We describe the rationale for this parameter in the documents linked from [this cell](#) in our cost-effectiveness analysis.

⁵ We describe our informal Bayesian approach [here](#) and [here](#). The rationale for our 13% replicability adjustment for deworming is in the documents linked from [this cell](#).

We think it's possible these changes reflect true declines in effect over time and that we should account for this possibility in our CEA. We do this by (i) putting some weight on these providing estimates of different effects over time, (ii) updating from a prior that effects are constant over time, and (iii) applying separate replicability adjustments for each survey round and using effects from KLPS-2 to KLPS-4 to extrapolate declines 40 years out.

Weight on decay

When we look at evidence like this, we typically favor pooled results when there is no a priori reason to believe effects differ over time, across geographies, etc. (e.g., a meta-analysis of RCTs for a malaria prevention program). In cases where there's more reason to believe the effects vary across time or geographies, we're more likely to focus on "sub-group" results, rather than pooled effects. In either case, this is often a subjective assessment.

In this case, we're uncertain about whether to pool results or not and think there are reasons for and against putting more weight on decline in effects over time. As a result, we put 50% weight on the surveys capturing noisy estimates of the same effect and 50% weight on surveys capturing true changes in effects on earnings and consumption over time.

Reasons for putting more weight on effects varying over time:

- The point estimates we have from KLPS-2, KLPS-3, and KLPS-4 show a [decline over time](#).
- There are plausible stories for why effects would decline. For example, it's possible individuals in the control group are catching up to individuals who were dewormed due to broader trends in the economy. This is speculative, however, and we haven't looked into drivers of changes over time.

Reasons to put less weight on effects varying over time:

- The evidence for decline comes from three noisy estimates of income and two noisy estimates from consumption (By noisy, we mean the estimates have wide, overlapping confidence intervals).⁶ It's possible that the observed decline is due to chance.
- There are differences in how data were collected across rounds that limit comparability of effects over time and that may drive the observed decline over time:
 - In KLPS-2 a lot of the sample was still in school,⁷ so it might be incorrect to look at that round on its own and think of it as representative of the full sample.

⁶ [Hamory et al. 2021](#), Appendix, Fig. S3.

⁷ "It is worth noting that one quarter of both the treatment and control groups are still in school by the time of the survey (Table II), and labor market outcomes are less meaningful for this group." [Baird et al. 2016](#), IV.C. "Impact on Labor Hours and Occupation," paragraph 1.

- Higher In earnings effects from KLPS-2 to KLPS-3 are driven by lower control group earnings in KLPS-2 (\$330 vs. \$1165).⁸ In KLPS-3, researchers started measuring farming profits in addition to other forms of earnings,⁹ so part of the apparent increase in control group earnings from KLPS-2 to KLPS-3 is likely driven by a change in measurement, not real standards of living or catch-up growth.
- The big increase in control group earnings from KLPS-3 to KLPS-4 (\$1165 to \$2133)¹⁰ is especially surprising and potentially questionable because there doesn't appear to be any change in control group average consumption from KLPS-3 to KLPS-4. If anything, it looks like there's a decline (\$2878 vs. \$2044),¹¹ though those measures have wide confidence intervals.
- There is a [decline in the effect on In consumption](#) from KLPS-3 to KLPS-4. However, the consumption effect in KLPS-3 was in a small sample¹² and unexpectedly large. [We funded KLPS-4](#) and the consumption effects came more in line with what we expected, which is why we didn't see it as a “decline.”
- We conducted a shallow literature review of studies of the effect of health interventions during childhood on adult income, and we found a similar number of studies finding a decline in effect as an increase in effect over time. If we found strong evidence that this type of program yields declining effects over time, we would put more weight on this story (see [below](#) for more detail).
- It seems plausible that income effects would be constant over time or could compound over time. For example, adults who were dewormed as children and see greater cognitive or educational gains may be less likely to enter sectors like agriculture, which we believe may have flatter earnings trajectories, or be more likely to move to cities, where opportunities for wage growth may be higher. However, these stories are also speculative.

We're uncertain about the appropriate weight to put on the interpretation that income effects are different (and declining) over time, and this is a key judgment call in our analysis.

⁸ [Hamory et al. 2021](#), Appendix, Fig. S3. “Deworming Treatment Effects by Survey Round, B. Annual Individual Earnings.”

⁹ “Annual individual earnings are calculated as the sum of wage employment across all jobs; nonagricultural self-employment profit across all business; and individual farming profit, defined as net profit generated from noncrop and crop farming activities for which the respondent provided all reported household labor hours and was the main decision maker within the last 12 mo. Wage earnings and self-employment profits were collected in KLPS-2, KLPS-3, and KLPS-4; agricultural profits were collected in KLPS-3 and KLPS-4.” [Hamory et al. 2021](#), Table 2.

¹⁰ [Hamory et al. 2021](#), Appendix, Fig. S3. “Deworming Treatment Effects by Survey Round, B. Annual Individual Earnings.”

¹¹ [Hamory et al. 2021](#), Appendix, Fig. S3. “Deworming Treatment Effects by Survey Round, A. Annual Per-Capita Consumption.”

¹² “The measurement of economic outcomes was also improved: KLPS round 4 (KLPS-4) incorporates a detailed consumption expenditure questionnaire (modeled on the World Bank Living Standards Measurement Survey; see ref. 32) for all respondents, and round 3 collected this for a representative subsample.” [Hamory et al. 2021](#), Introduction, paragraph 5.

Prior for decay

A key assumption is that we're updating from a prior that the effects on increased income are 1% and constant over 40 years. If we had reason to believe instead that effects should decay, based on evidence from similar interventions, then we'd be updating from a prior of decay and include steeper decay. We would also put more weight on the interpretation that the different estimates for the effect of deworming over time are capturing true differences and less on the interpretation that these are noisy estimates of the same effect.

In order to assess whether the impact of deworming on income increases, decreases, or remains the same over the lifecourse of those receiving deworming treatment as children, we carried out a shallow literature review and consulted with experts and GiveWell researchers regarding studies of childhood interventions with multiple adult follow-ups. We looked for studies that examined long-term effects of improvements in early-life health (e.g., weight/height), cognition, and education, which we think are some of the plausible mechanisms through which deworming leads to impacts on later-life income.

We found 10 longitudinal studies with at least two adult follow-ups from a number of countries examining the impact of a range of childhood interventions or conditions (see [this table](#)), in addition to the deworming study ([Hamory et al. 2021](#)).

Of those 10 studies, [3 found decreasing effects on income](#), [2 found increasing effects](#), and [5 found mixed effects](#) (either similar effects across time periods, different patterns across males and females, or increases and then decreases over the life cycle).

Based on this, we think it makes sense to continue to assume as a prior that income effects would be constant over time. I have low confidence in these estimates, though, and it's possible further work could lead to a different conclusion. Specific areas of uncertainty and areas for further investigation are:

- We did not do a deep review of studies. We did a quick scan to see if authors reported changes in effects over time. As a result, there might be some nuances of comparing across time we've missed.
- It's possible we've missed some relevant studies altogether.
- We have not tried to formally combine these to get point estimates over time or attempted to weight studies based on relevance, study quality, etc.
- We are combining studies that may have little ability to inform what we'd expect from deworming (twin studies, childcare programs, etc.).
- It could be possible to re-assess other studies measuring long-term benefits of early childhood health interventions. When we set our prior, we excluded studies that did not report separate effects on income at different time periods. We guess that for several of these studies, it would be possible to re-analyze the primary data and create estimates of the effect on income at different time periods.

- We could poll experts working in this field to get their best guess on the extent to which effects would fade over time or not.
- We're also aware that there is an additional survey underway (KLPS-5) that will collect detailed consumption data. We expect to be able to update based on the results of that study as well.

Replicability adjustment for each survey

We use a replicability adjustment in our deworming CEA to capture our best guess at the portion of the income effects of deworming found in the Kenya RCT that would be found if a perfect experiment could be run again under the same conditions. To create this adjustment, we use a broadly Bayesian framework.¹³ Our “prior” in this context is our best guess at what we would have expected the effect size of deworming on developmental effects to be in absence of results from the Kenya RCT. We then update our prior using the Kenya RCT and our views on the strengths and limitations of the evidence base.

To incorporate decay into our estimates, we apply separate replicability adjustments for each follow-up survey from the Kenya RCT (KLPS-2, KLPS-3, and KLPS-4). Under each story (different estimates over time vs. noisy estimates of the same effect), we update from a [prior of 1%](#) impact on consumption over time. I updated replicability adjustments for each of the estimates (10 years, 15 years, 20 years) by running the same replicability adjustment calculations for each year. In the case where we interpret these as different estimates over time, I follow a similar approach to our current CEA but update separately for each time period.

Our current approach in the [deworming CEA](#):

- We interpret the 3 effects (from the 10-year, 15-year, and 20-year follow-ups) from KLPS-2, KLPS-3, and KLPS-4 as three noisy estimates of the same effect.
- We currently apply a [13%](#) replicability adjustment to an estimated average effect of [0.109](#) on log income/consumption. This is based on (1) updating from a skeptical prior based on mechanisms analysis, (2) updating from a skeptical prior based on an informal Bayesian update, and (3) updating based on an informal qualitative case. Write-up [here](#). Spreadsheet [here](#).
- This is intended to capture both our uncertainty about program impact and our prior that the true effects of deworming on later-life income are much smaller than what is found in this study.
- Our best guess is a [~1.4%](#) increase in income/consumption across 40 years.¹⁴

The alternative approach (which views KLPS-2, KLPS-3, and KLPS-4 as capturing different income effects and so allows there to be decay):

¹³ See [this blog post](#) for further discussion of GiveWell's approach to using broadly Bayesian frameworks in our analyses.

¹⁴ 1.4% equals [0.109](#) treatment effect * [13%](#) replicability adjustment.

- We're [re-doing the replicability adjustment calculations](#) but separately for each survey/time period (KLPS-2/10 years, KLPS-3/15 years, KLPS-4/20 years).
- We set the same prior as before (1% effect over 40 years) and update from that at 10 years, 15 years, and 20 years.
- We end up with a 7% adjustment on the 0.234 effect on log income/consumption in KLPS-2, 8% adjustment on the 0.185 effect in KLPS-3, and 19% adjustment on the .066 effect in KLPS-4. The calculations are in [this spreadsheet](#).
- Our best guess is then [~1.6%](#), [~1.5%](#) and [~1.3%](#) effects in years 10, 15, and 20. I extrapolate to year 40 by taking the exponential trend from year 10 to year 20 in [this spreadsheet](#).
- I don't feel very confident in the quantitative estimates of replicability, but intuitively, it feels right that, if we viewed these as separate estimates over time, (1) we'd update toward an effect size for deworming on income/consumption higher than ~1.4% (our current best guess) in year 10 and year 15 ([where estimated effects are larger than the current average we use](#)) and lower in [year 20](#) and (2) the gradient wouldn't be that steep, since the effects on income over time are noisy and may be capturing different measures of income and consumption over time,¹⁵ which means we're not that responsive to fluctuations over time.

Like our current replicability adjustments, these estimates hinge on judgment calls and assumptions.

- Our priors for the effect of deworming are based on a rough analysis that includes several subjective assessments. These are described [here](#) and [here](#). Because I am extending this approach (by updating our prior separately for the 10-, 15-, and 20-year data), the decay model is subject to these same limitations.
- Hamory et al. (2021) do not report estimates of ln earnings and consumption by round, so we have to approximate effects in ln and their standard errors. (See calculations [here](#).)
- We're unsure about how to extrapolate effects beyond the 20-year follow-up. An alternative approach would be to assume any declines from 10- to 20-year follow-up begin to level off, which would weaken the adjustment.
- In the scenario where we assume KLPS 2-4 are estimating separate effects, we're assuming the estimates are totally independent. Even if we thought these were measuring decay over time, that assumption seems incorrect, since we're tracking the same kids over time. This seems like it would increase the effect size across rounds.

More broadly, there may be alternative approaches to updating from priors on both the average effect of deworming and decay over time that are more accurate. We've chosen to model decay by (i) specifying a prior on the effects of deworming on income over time and (ii) updating from this prior by putting some weight on the RCT in Kenya finding decay in effects over time and some weight on the RCT capturing noisy estimates of the same effect over time. There may be better or more formal approaches to model decay (e.g., by putting priors on the initial effect of

¹⁵ See discussion above.

deworming and a prior on decay, then updating both based on the KLPS surveys). Ultimately, we chose the current approach because it seems like the most straightforward and most consistent with what we're currently doing, but it's possible alternative approaches are better.

Bottom line adjustment factor

Our best guess is that we should apply a [-10%](#) adjustment due to the possibility of decay in effects over time.

In the model where we assume KLPS 2-4 provide noisy estimates of the same effect, we estimate an average effect of deworming of [0.109](#) on ln income. When we update from our skeptical prior, our best guess is [~1.4%](#) over 40 years for a net present value of [0.115](#).

In the model where we assume KLPS 2-4 provide different estimates over time, we estimate an effect of [0.23](#), [0.19](#), and [0.07](#) on ln income at 10-years, 15-years, and 20-years post deworming. When we update from our skeptical prior, our best guess is [~1.6%](#), [~1.5%](#), and [1.3%](#) at years 10, 15, and 20 and a net present value of [0.093](#) during the full time period.

We put 50% weight on each of these interpretations, which lowers the total effect by [-10%](#) (relative to putting 100% weight on KLPS 2-4 capturing noisy estimates of the same effect).

Sources

Baird et al. 2016	https://doi.org/10.1093%2Fqje%2Fqjw022
GiveWell, "2023 cost-effectiveness analysis"	https://docs.google.com/spreadsheets/d/10JFJaWnFAEKmsv5XjXqGqEoMUx0eM7x3WYwuvC7FRw/edit#gid=472531943
GiveWell, "Context on deworming replicability adjustment (2020)"	https://docs.google.com/document/d/1-F5sZBq6FD6E73SWkKFhwMR9gCdKUCTfp9dOe0l-1vw/edit

GiveWell, "Deworming decay adjustment: deworming effect calculation (2023)"	https://docs.google.com/spreadsheets/d/1iUcljfudwQlPOftbG_e5axbiAfRr15ie7rvfTiiDFvU/edit#gid=1321957472
GiveWell, "Deworming decay adjustment: KLPS 4 Deworming Effect Size Parameter Update (2023)"	https://docs.google.com/spreadsheets/d/1bbZWTjklQ5hc2i4zynCR6gg2TqR0x3HCgC-ssK6i0FI/edit#gid=1667455426
GiveWell, "Deworming decay adjustment: replicability adjustment (2023)"	https://docs.google.com/spreadsheets/d/1u6kDrFbns-2_M46G_POro09RcKQZ0TSy8gg-IXK6Z1o/edit#gid=2002315610
GiveWell, "Deworming decay adjustment: replicability adjustment (informal Bayesian analysis, 2023)"	https://docs.google.com/spreadsheets/d/1kOh43pku33n7AQyAv6X43Z9bTUThcE43ZZcOeD-YUoo/edit#gid=251688210
GiveWell, "Deworming Effect Size Parameter Update - KLPS 4 Results 11.07.19"	https://docs.google.com/spreadsheets/d/1MNEPqRhIndfpeJT3LxCk1Bvrn-N81n0PlvsXRjc3fb4/edit#gid=

	0
GiveWell, "Deworming replicability adjustment (2020)"	https://docs.google.com/document/d/1PZfYXegWco0qrmQnQBjclZeq4uUdWfpE5xdBYx0cAEU/edit
GiveWell, "Deworming replicability adjustment 2019"	https://docs.google.com/spreadsheets/d/1ZvX6XI5AKxTYQJbyxlEkmuf6LyYlt18kyfeqRGs-aaM/edit#gid=0
GiveWell, "Long-term effects literature review"	https://docs.google.com/spreadsheets/d/1n1fRU77jvxFlkiHF4zoQn3l6n3KcpiGfLih_W_JniM/edit#gid=0
GiveWell, "UC Berkeley — KLPS-4 Survey"	https://www.givewell.org/research/incubation-grants/uc-berkeley/april-2017-grant
GiveWell, "Why we can't take expected value estimates literally (even when they're unbiased)," 2011	https://blog.givewell.org/2011/08/18/why-we-cant-take-expected-value-estimates-literally-even-when-theyre-unbiased/
Hamory et al. 2021	https://doi.org/10.1073/pnas.20231851

	18
McGuire, Dupret, and Plant, "Deworming and decay: replicating GiveWell's cost-effectiveness analysis," 2022	https://web.archive.org/web/20230221162055/https://forum.effectivealtruism.org/posts/MKiqGvijAXfcBHCYJ/deworming-and-decay-replicating-give-well-s-cost