BENEFICIARY PREFERENCES

FINDINGS FROM GHANA AND KENYA

7th November, 2019

IDinsight
DATA. DECISIONS. DEVELOPMENT.

Authors

Alice Redfern: alice.redfern@IDinsight.org

Martin Gould: martin.gould@IDinsight.org

Maryanne Chege: maryanne.chege@IDinsight.org

Sindy Li: sindy.li@IDinsight.org

Felipe Acero Garay: felipe.acerogaray@idinsight.org

William Slotznick: william.slotznick@IDinsight.org

About IDinsight

IDinsight uses data and evidence to help leaders combat poverty worldwide. Our collaborations deploy a large analytical toolkit to help clients design better policies, rigorously test what works, and use evidence to implement effectively at scale. We place special emphasis on using the right tool for the right question, and tailor our rigorous methods to the real-world constraints of decision-makers.

IDinsight works with governments, foundations, NGOs, multilaterals and businesses across Africa and Asia. We work in all major sectors including health, education, agriculture, governance, digital ID, financial access, and sanitation.

We have offices in Bengaluru, Dakar, Johannesburg, Lusaka, Manila, Nairobi, New Delhi, San Francisco, and Washington, DC. Visit www.IDinsight.org and follow on Twitter @IDinsight to learn more.

TABLE OF CONTENTS

Executive Summary..... 4

Results Overview.....6

Introduction.....15

Section 1: Value of Increasing Consumption Relative to Saving Lives.....18

Section 2: Relative Values of Individuals Under 5 Vs. 5 and Older.....29

Section 3: Moral Reasonings.....35

Section 4: Subjective Wellbeing.....39

Section 5: Economic and Emotional Impact of Death.....44

Conclusion.....51

References.....54

Appendix 1: Sample Approach and Demographics.....57

Appendix 2: Individual Perspective: Value of Statistical Life (VSL).....60

Appendix 3: Community Perspective: Monetary Value of Life.....71

Appendix 4: Relative Value of Age Groups.....82

Appendix 5: Moral Reasonings Data.....86

Appendix 6: Subjective Wellbeing - Additional Data.....92

Appendix 7: Economic and Emotional Burden of Death – Additional Data.....96

Appendix 8: Aggregated Results from Primary Methods.....102

Appendix 9: Findings from Secondary Methods 105

Appendix 10: Study Limitations.....113

EXECUTIVE SUMMARY

CONTEXT

International development leaders frequently make complex resource allocation decisions that require weighing trade-offs between different types of good outcomes. For example, given limited resources, which should be prioritized: a program that increases household income or one that saves lives?

When comparing diverse charities, GiveWell makes these decisions transparent by asking staff members to provide their 'moral weights.'[1] These judgments are based on philosophical reasoning, intuition and data on beneficiary lives, and extrapolation of preferences from studies of less relevant populations. Prior to this study, there was a clear lack of data on how potential beneficiaries of such interventions trade-off between different outcomes. This study represents a step to fill this gap for strategic international development decision-making.

APPROACH

We surveyed over 1,800 low-income individuals across four diverse regions in Ghana and Kenya. Three main methods[2] were used to capture how respondents trade-off between averting deaths of individuals of different ages and increasing consumption:

1. We asked individuals for their willingness-to-pay (WTP) to reduce the risk of death for themselves and their children.[3]
2. We asked respondents to take the perspective of a decision maker in their community and choose between programs that:
    a. Save lives of different ages;
    b. Save lives and provide cash transfers.

We also collected qualitative data on beneficiaries' reasoning when making these trade-offs, and data on beneficiaries' lives that can be used to inform GiveWell staff's moral weights.[4]

[1] Moral weights are subjective numerical values that weight the trade-off between deaths averted and gains in household consumption. Once staff members have provided their individual moral weights, the median across the group is taken to inform their cost-effectiveness model. https://www.givewell.org/how-we-work/our-criteria/cost-effectiveness/comparing-moral-weights

[2] The three methods were selected after extensive piloting of over 15 different approaches. They were selected based on: 1) their relevance to GiveWell moral weights; and 2) the reliability of the methods to collect accurate data.

[3] Value of Statistical Life (VSL), an approach commonly used by governments to inform cost-effectiveness analysis.

[4] This includes: 1) primary data on the subjective well-being of our respondents, and how this correlates with individual characteristics; and 2) primary qualitative data on the economic and emotional impact of household deaths; and 3) secondary data analysis of the relative economic contribution of different household members in a typical beneficiary population.

FINDINGS

We found that:

- Respondents place a higher value on averting a death[5] than predicted by most extrapolations from studies in high income countries (HICs).
    - Our central estimate of value placed on averting death for individuals 5 and older was $40,721, which is 1.7 times higher than the current GiveWell staff median.
- Respondents consistently value the lives of individuals under 5 higher than individuals 5 and older, which is consistent with HIC studies but contrary to median GiveWell moral weights.
    - Our central estimate of value placed on averting death for individuals under 5 was $65,906, which is 4.9 times higher than the current GiveWell staff median

Qualitative data suggests these high valuations are driven by a large proportion of individuals making two arguments. The first argument asserts the importance of accounting for the potential held by all individuals to achieve high economic and social value over their life-course. A second common argument is that life holds inherent value and therefore no amount of money is sufficient to forego the chance to save a life.

IMPLICATIONS

For GiveWell, incorporating the preferences captured in this study and described above would result in:

1. Placing a higher value on averting deaths relative to doubling consumption,
2. Placing a higher value on averting the death of individuals under 5 than individuals 5 and older.

This would lead to a higher relative cost-effectiveness of charities whose good is achieved primarily by averting the death of young children (e.g. Helen Keller International, Malaria Consortium, and Against Malaria Foundation etc.).[6]

Beyond GiveWell, this study demonstrates that it is possible to inform the complex, moral trade-offs faced in development by capturing the preferences of the people affected by these decisions. It also represents a substantial addition to existing literature on individual cross-outcome preferences, in which low income populations and particularly those in Sub-Saharan Africa, have previously been severely underrepresented.

In a development sector which is generally reluctant to rigorously compare different types of outcomes, we believe this study could make the use of portfolio-level cost effectiveness analysis more appealing. We encourage additional research to further develop understanding of beneficiary preferences across program areas and from different populations.

[5] This includes death and risk of death for themselves, and individuals in their family and their community.

[6] If the moral weights were changed to fully reflect the quantitative results of this study, the ranking of current GiveWell top charities would shift substantially and Helen Keller International, Malaria Consortium, and Against Malaria Foundation would be the most cost-effective charities (with charities such as Sightsavers and GiveDirectly becoming less cost-effective). It would also make it substantially easier for other charities whose main outcome is saving lives, particularly of children under 5, to reach and surpass GiveWell's cost-effectiveness threshold (2-3x as cost-effective as cash).

RESULTS OVERVIEW

APPROACH

Our three primary methods (see Table 1) were designed to inform two components of GiveWell's moral weights:

- **Component 1:** the value assigned to averting the death of an individual relative to doubling consumption for one person for one year.
- **Component 2:** the value assigned to averting the death of an individual under 5 relative to an individual 5 and older.

Our methods ask questions from two perspectives: individual (regarding respondents' own outcomes), and community (regarding outcomes affecting poor households within respondents' own community).

Table 1. Summary of three primary methods.

Method	Component	Perspective
Value of Statistical Life (VSL): Contingent valuation method eliciting willingness to pay for small reductions of mortality risk for self/child.	Data to inform both 1 & 2	**Individual:** Respondents are asked for preferences over outcomes for themselves and their own children. This approach was designed to align with the VSL literature.
Monetary value of life: Choice experiment measuring preferences for programs which save lives, and programs which increase household consumption, in their community.	Data to inform 2 only	**Community:** Respondents are asked to take the perspective of a decision maker and make trade-offs on allocating resources within their community. This approach was designed to mirror the trade-offs faced by a resource allocator, such as GiveWell.
Relative value of age groups: Choice experiment measuring preferences for saving lives of people of different age groups.	Data to inform 1 only	

We also used several secondary methods to capture data on the following:

1. The **moral reasonings** used by beneficiaries when making trade-offs related to Component 1 and 2.
2. The **subjective wellbeing** of beneficiaries (as measured by self-reported life satisfaction), including how this correlates with different individual characteristics.
3. Other **information about beneficiary lives**, including primary data on the indirect effects of death (economic and emotional) and secondary data analysis on economic contribution by age.

These methods were refined over the course of 12 months of piloting, including consultation with experts in the field of VSL and preference elicitation. The three primary methods were chosen as they are directly relevant to GiveWell moral weights, and we had reasonable confidence they would

produce reliable and internally valid data at scale.[7] For smaller subsamples, we varied the content or framing of the primary methods to collect data on additional areas of interest.[8]

We surveyed respondents from low-income households (see Appendix 1 for a description of our sampling approach). In addition to 1,846 'poor' respondents, we also surveyed 246 'wealthy' respondents from wealthier households in the same communities. These 'wealthy' are not included in the main estimates, but are included in certain analyses.[9]

FINDINGS

COMPONENT 1: THE VALUE ASSIGNED TO AVERTING THE DEATH OF AN INDIVIDUAL RELATIVE TO DOUBLING CONSUMPTION FOR ONE PERSON FOR ONE YEAR

Respondents placed substantially higher weight on averting death relative to doubling consumption than predicted by most of the standard methods of extrapolation from high income countries' (HIC) data, and than GiveWell's current aggregate moral weights (see Table 2).

The qualitative analysis indicates this is driven by a large proportion of the sample who place high weight on the idea that life holds inherent value and, therefore, no amount of money is sufficient to forego the chance to save a life. Respondents also frequently mentioned the need to protect the potential held by all individuals to achieve high economic and social value over their life-course, leading them to place an extremely high value on life, particularly for young children.

Value of life results are consistently higher in the community perspective compared to the individual perspective – this likely reflects an underlying difference in what each method captures. As the individual perspective approach is based on WTP, even with integrated steps to overcome a likely liquidity constraint,[10] respondents are still constrained in what they are willing to pay by their income (i.e. what they are able to pay). In contrast, in the community perspective, respondents weight trade-offs as a decision-maker, and so their own income should not constrain their decisions. This and other differences[11] in what each method captures may affect how GiveWell staff members consider using these results (as explored further below in).

[7] See the IDinsight report, "Beneficiary Preferences: Proposal for Scale-up 2019", for more details.
[8] This includes two alternate framings of our choice experiment, one in which a comparison is made between life-saving programs and education programs and one in which the trade-off between averting deaths and increasing consumption is presented more directly. The results are included in Appendix 9.
[9] Specifically, they are only included in the analyses of the relationship between wealth level (consumption per capita) and VSL and life satisfaction, as well as subgroup analyses for the choice experiments.
[10] 'Liquidity constraint' refers to a constraint on an individual's ability to pay for something, due to an inability to borrow caused by imperfect capital markets. To overcome this constraint, we ask respondents for their WTP in small instalments over 10 years such that each individual hypothetical payment is much less influenced by any liquidity issues.
[11] There are other reasons why results may vary between the two methods, including that (a) the individual perspective asks respondents to value their life or their child's life, whereas the community perspective asks respondents to value community members; (b) the individual perspective involves reducing mortality risk by a small amount, whereas the community perspective involves saving lives with certainty.

Table 2. Comparison of derived value (USD) of individuals under 5 and those 5 and older, between previous estimates, and our study results.

	Derived value (USD) of an individual under 5	Derived value (USD) of an individual 5 and older
Reference 1: Current median of staff moral weights[12]	$13,505 ($5,720 -$21,056)	$24,406 ($8,580 - $45,660)
Reference 2: Current GiveWell 'conventional' figure[13]	$26,119	$21,286
Reference 3: Literature prior, based on best experimental evidence to date[14]	N/A	$27,765 ($1,876 - $41,091)
New evidence 1: Individual perspective - WTP for mortality risk reduction for self and own child (VSL)	$40,763	$33,798
New evidence 2: Community perspective – trade-off between life-saving and income increasing programs for community	$91,049	$47,645[15]

While our primary methods indicate beneficiaries place a higher value on life relative to increasing consumption compared to GiveWell's current moral weights, our other findings demonstrate the importance of increasing consumption levels for beneficiary self-reported wellbeing. The subjective wellbeing of our respondents is on average low.[16] Doubling an individuals' per capita consumption is associated with being 0.4 steps higher on the life satisfaction ladder, and when asked what factors would increase their life satisfaction a large proportion mentioned money or material goods (see Section 4 for more information).

[12] This figure is taken from the 'aggregate' column in GiveWell's current cost-effectiveness model. The range represents the highest and lowest staff moral weights. We convert between moral weights and USD assuming that one unit of moral weights is equivalent to the value of doubling consumption for one individual for one year as per the current model ($286). For more discussion of this conversion see Appendix 8.

[13] This figure is taken from the 'conventional column' in GiveWell's current cost-effectiveness model. This approach values a DALY at 2.5 times beneficiaries' annual consumption per capita and assumes averting an adult death is equivalent to averting 30 DALYs, and averting a child death 37 DALYs.

[14] The median value is based on the extrapolation from median VSL for HIC to the target population assuming an income elasticity of 1.5, following the recommendations set by Robinson et al. (2019). The lower end of the range assumes an income elasticity of 2.0, and the upper end assumes an income elasticity of 1.0.

[15] Note that this value is derived by applying the ratio derived from our choice experiment on the relative values of age groups (1.9) to the value derived for individuals under 5 from our choice experiment on the monetary value of life ($91,049). For simplicity, we assume independence between the two estimates and use simulations to obtain the mean and standard deviation of the distribution of the product.

[16] Self-reported life satisfaction among poor individuals was 2.8 on a ladder between 0 ("The worst possible life for you") and 10 ("The best possible life for you").

COMPONENT 2: THE VALUE ASSIGNED TO AVERTING THE DEATH OF AN INDIVIDUAL UNDER 5 RELATIVE TO AN INDIVIDUAL 5 AND OLDER

Respondents consistently place higher weight on individuals under 5 than individuals 5 and older – these results are within the range of predicted values from the literature (see Table 3). This contrasts with the lower weight placed on individuals under 5 in GiveWell's current moral weights.

As noted above, our results are largely driven by respondents placing high value on the potential for young children to achieve an extremely high economic or social value over their life course. Respondents also frequently argued that young children should be saved out of fairness, or out of a responsibility that adults hold towards the young and vulnerable.

Table 3. Comparison of derived value (USD) of the relative value of individuals under 5 to those 5 and older, between previous estimates and our study results.

	Value of an individual under 5 relative to an individual 5 and older
Reference 1: Current median of staff moral weights	0.55 (0.32-1.43)
Reference 2: Current GiveWell conventional figure	1.23 (no range)[17]
Reference 3: Literature prior, based on best experimental evidence to date[18]	>1.5 (0.6-2.9)
New evidence 1: Individual perspective – ratio of WTP for own child under 5 to that for self or child 5 and older	1.25
New evidence 2: Community perspective – trade-off between programs saving individuals under 5 vs 5 and older	1.9 excluding > 40-year-olds 3.7 including > 40-year-olds

The relative value placed on individuals under 5 drastically increases when results related to individuals over 40 are included in our analysis (the ratio increases from 1.9 to 3.7).[19] As the value placed on individuals over 40 was so low (and therefore has such a large impact on the results), and as GiveWell top charities do not typically serve this age group, we suggest that the ratio excluding individuals over 40 is of greater relevance to moral weights.

[17] This assumes that averting a death of an adult is equivalent to averting 37 DALYs, and averting that of a child 30 DALYs.
[18] This range is taken from Robinson et al. (2019) recent summary of HIC literature on this ratio.
[19] Our results found a negative valuation of individuals over 40, suggesting that it would be preferable to respondents to have fewer older individuals within the community (i.e. in a choice between 100 individuals over 40 and 200 individuals over 40, respondents would choose the former). However, we are less confident about the estimate on people over 40 than those on other age groups. See Section 2 for a discussion of our interpretation of the negative estimate.

INCORPORATION INTO MORAL WEIGHTS

Note: Developing a single estimate of moral weights estimate from the results of this study is not straightforward and requires a number of significant assumptions. Here we present our best central estimate of preferences and explore how it might be incorporated into moral weights. However, we recommend reading the full report to understand the results in totality and to make an informed choice on how to aggregate and use the results.

In order to form a single central estimate of beneficiary preferences that can be incorporated into GiveWell's CEA model, we aggregate results across our main approaches by weighting the individual and community perspectives equally, then convert our value from USD into moral weights (for results, see Table 4). The full process for aggregating and converting results is outlined in Appendix 8.

Fully incorporating beneficiary preferences[20] into moral weights would:

- **Increase by a factor of 1.7 (for individuals 5 and older) and by a factor of 4.9 (for individuals under 5)** the value placed on averting death relative to doubling consumption.
- **Increase the value** placed on individuals under 5 relative to individuals 5 and older **from 0.6 to 1.7.**

Table 4. Aggregated beneficiary preferences, in comparison to the current median of GiveWell staff moral weights.

	Derived value (USD)		Moral Weights	
Value assigned to…	Aggregated beneficiary results	Current GiveWell median	Aggregated beneficiary results	Current GiveWell median
Averting the death of an individual under 5	$65,906	$13,505	230	47
Averting the death of an individual 5 and older	$40,721	$24,606	142	85
Doubling consumption for one person for one year			1	1

GiveWell staff members may wish to use the beneficiary preferences results in combination with other considerations when developing their moral weights. Below we set out an approach for doing this.

1. **Decide how much weight to place on beneficiary preferences, relative to other factors that may inform moral weights.**

There are several factors beyond beneficiary preferences that can inform moral weights, ranging from philosophical deliberations to empirical data on beneficiaries' lives.[21] Staff members may choose to defer completely to beneficiary preferences, to only use other factors, or to combine beneficiary preferences with other factors (giving a relative weighting to each).

[20] By fully incorporating we specifically refer to giving 100% weight to our aggregated results, such that moral weights solely reflect the central estimate of beneficiary preferences.

[21] For more detail on the range of factors that are considered, see the accompanying moral weights framework documentation.

When deciding on how to weight beneficiary preferences relative to these other factors, it is useful to consider the extent to which these other factors are considered by beneficiaries and therefore reflected in the above results. For example, a factor some GiveWell staff members consider is the indirect effects of a death on the family and friends of that individual.[22] Many respondents reported incorporating the indirect effects of death in their responses to our three primary methods – so, in effect, the consideration of indirect effects of death is included in our quantitative results in Table 4.[23]

2. Decide how much confidence to place in the methods and data collection which produced this study's estimates of beneficiary preferences

It is possible to value beneficiary preferences, but also believe that the methods used in this study or the way they have been applied do not capture the preferences of GiveWell beneficiaries. If this is the case, staff members may want to adjust their relative weighting of beneficiary preferences compared to other factors, or may want to lean more heavily on the prior literature.

Some considerations for this include:

- *Confidence in the reliability of the methods:* We developed and tested the study methods over 12 months drawing on the latest literature and interviews with experts in the field. Our methods rely on stated preferences design, which has well documented challenges. However, with few opportunities for revealed preference studies in this field, we believe the methods we used are the most appropriate and reliable to elicit beneficiary preferences.
- *Confidence in the reliability of our application of the methods:* Respondent understanding of our approaches was good, and we passed key validity tests across all methods.[24] This leads us to have relatively high confidence in the reliability of our results. On the other hand, compared to all the studies completed in HICs from which data can be aggregated and extrapolated, our sample size is small.
- *Relevance of the study population to the decision*: Prior to our study, there was limited data related to the life-consumption trade-off from LMICs.[25] It was common for researchers to take data from HICs and apply a linear extrapolation based on income to estimate values for LMIC populations. By capturing preferences from a sample of respondents similar to typical GiveWell beneficiaries, the results of this study offer a substantial improvement on the best estimates previously available.[26]

3. Choose whether to adjust the process we have followed to compile our single aggregated results

Our aggregation process makes several assumptions that may not be held by GiveWell staff members. Most importantly, we assume that the individual and community perspectives are of equal relevance

[22] For example, the emotional toll of the death, and the financial impact on a family of losing an individual whose income supports the family.
[23] See Section 4: Moral Reasonings and Appendix 5: Moral Reasonings Data, for more information on the range and frequency of factors considered by beneficiaries when weighting trade-offs.
[24] For more information on the evidence for and against the reliability of our approaches see Appendix 2-4.
[25] Some existing studies from low and lower-middle income countries include Kremer et al. (2011), Mahmud (2011), Léon and Miguel (2017), and Shrestha (2017).
[26] For more information about our sample, see Appendix 1.

to moral weights Table 5 summarizes the attributes of these approaches to support staff members to decide on the relevance of each. In addition, we make assumptions about which is the most appropriate central estimate to use, how to reflect relative confidence in the two approaches in our weighting, and how to convert between USD and the moral weights unit. In Appendix 8, we explain each of our assumptions in turn, and demonstrate how the results vary if differing assumptions are made.

Table 5. Comparison of individual and community perspective, to inform relevance weighting.

	Individual Perspective	Community Perspective
Description	Respondents are asked for preferences related to themselves and their own children. Specifically, we capture VSL – an individual's WTP for small mortality risk reductions.	Respondents are asked to take the perspective of a decision maker and make trade-offs about how to allocate resources within their community. Our two choice experiments ask respondents to decide between programs that save lives (with certainty) and provide cash transfers, and programs that save lives of different ages (with certainty).
Comparability to literature	Approach is commonly used in the literature.	Approach used in the literature, but less common than VSL.[27]
Approach used by other actors	Results from this approach are frequently used by governments.	Lack of examples of government or other actors using results from this approach.[28]
Comparability to moral weights	VSL questions are from the individual perspective rather than the decision maker / moral weights perspective.	Question framing is directly comparable to moral weights.
Method constraints	Applies a WTP methodology, and so is inherently income constrained.[29]	Social desirability bias may lead to respondents stating they always prefer life-saving programs.[30]
Impact of giving greater weight to this perspective	As the median VSL result is lower than the community perspective result, more weight on this method would result in a lower aggregated result than we present above.	As the median community perspective result is higher than the VSL result, more weight on this method would result in a higher aggregated result than we present above.

[27] Our method for estimating relative values of age groups has been used in at least three studies in Sweden and Bangladesh. Our method for estimating the monetary value of life (in terms of cash transfers) from a community perspective was designed specifically for this study.
[28] The ratio of child vs adult VSL is typically used by governments, rather than results from choice experiments.
[29] We believe we largely overcome the liquidity constraint by asking about payments in small instalments, but respondent income still creates a ceiling on what they are able to pay. This may limit the ability for respondents to exhibit really strong preferences for life saving interventions as we see in the community perspective.
[30] We believe we have limited this by using a framing that is focused on 'an additional death averted' instead of a framing comparing saving a life and cash transfers directly (see Appendix 9 for a comparison of the two approaches). However, some social desirability may remain.

IMPLICATIONS BEYOND GIVEWELL

This study demonstrates that it is possible to inform the complex, moral trade-offs faced in development by capturing the preferences of the people affected by these decisions.

This study represents a substantial addition to existing VSL literature in which low income populations, particularly those in Sub-Saharan Africa, have previously been severely underrepresented. The data produced can immediately inform benefit-cost analysis (BCA) used in policy decision making in related populations. Incorporating this study's results into BCA offers two main advantages:

1. *Incorporating data from the countries of interest increases confidence in BCA output, as the underlying assumptions better reflect the preferences of target populations. This may increase the likelihood of decision makers using these results.*

   Current recommendations for BCA in low income countries rely on extrapolation of VSL figures from studies in the US, due to the lack of reliable data from LMICs.[31] This raises the concern that the values do not represent the true preferences of the population served. This study provides a reliable reference point from low income households in Kenya and Ghana, a population that is more similar to those targeted by BCA in most LMICs.

2. *Our data provides a more granular look at how VSL varies within countries, specifically among very low-income populations.*

   Current recommendations for BCA provide estimates for country-level VSL,[32] but many BCA studies focus on interventions that target a specific section of the population. This study provides insight as to how VSL varies within country for Kenya and Ghana that can be used to inform assumptions where BCA results are required at the subnational level.[33]

Beyond the VSL community, these results may also be used by foundations, and individual donors, to directly inform their giving and allocation decisions. More work is required to explore how to make these results accessible to more audiences, and to understand how organizations beyond GiveWell can best use this data.

NEXT STEPS

This study fills a clear gap and can immediately inform GiveWell's moral weights and other BCA in low income countries. Our data focused on preferences related to one key trade-off, from a population of low-income respondents that are of high relevance to GiveWell's decision. However, a much broader range of trade-offs, affecting many diverse populations, are faced by decision makers. As such, we

[31] See Robinson et al (2019) guidance. While there are a small number of studies of VSL in LMICs, the majority have been based in China and South East Asia. These studies have informed recommendations about how to extrapolate from US figures to a lower income population based on GNI per capita. However, there is not enough primary data from LMICs of interest to come up with an estimate of VSL directly.

[32] Again, see Robinson et al (2019) guidance.

[33] For example, a recent study by IDSI conducted benefit-cost analysis to prioritise the use of different TB testing approaches (Wilkinson et al., 2019). The underlying VSL values used were those estimated for the whole of South Africa based on recommendations, despite the fact that TB disproportionately impacts the bottom wealth quintile. The outputs of this analysis are highly sensitive to this assumption.

believe this study should serve as the beginning of a larger push to capture and use beneficiary preferences across many areas to inform government and development sector resource allocation.

Some next steps to achieve this broader goal:

1. **Expand research agenda to capture:**
    a. **The preferences of populations across more diverse contexts**. Our sample was focused across 4 regions in Kenya and Ghana, and was not designed to be representative of these regions (as we were focussed on particularly poor households) or the population of these countries. There is a clear need for studies across more regions and countries to validate these results, and further explore variations in valuation between different populations.
    b. **Other types of preferences**. In line with GiveWell's priorities, this study focused on methods that directly capture the trade-off between increasing consumption and averting the deaths of individuals of different ages. However, there are many other beneficiary preferences that may be relevant to different decision makers. For example, our methods could be adapted to other non-health related outcomes that cannot be converted to dollars per lives saved. This could include preferences around the equality of distributions (i.e. should we target the poorest, or spread aid across larger populations), autonomy (i.e. should we prioritize public provision or interventions that increase beneficiary autonomy), and the relative value of health to other important outcomes (such as education or female empowerment).
2. **Work directly with different organisations to understand how others can incorporate preferences into their decision making.** This study was designed in partnership with GiveWell to inform the specific trade-offs they face, and the results presented to fit their decision-making system. By working with other organisation which face similar trade-offs we can better understand how others can use these results, and establish what additional data might be highest priority to collect.
3. **Develop tools that improve the accessibility of this data to a broader audience of policy makers and donors**. The resource allocation decisions addressed by this study, as well as the methods and analytical approach used, are complex. Simple tools that summarise the different ways to approach this problem, the best available literature, and the results of this study could make this decision-making process more accessible to a broader audience.

INTRODUCTION

CONTEXT

International development leaders frequently make complex resource allocation decisions that require weighting trade-offs between different types of good outcomes. For example, given limited resources, which should be prioritized: a program that increases household income or one that saves lives?

GiveWell conducts in-depth research on charities to identify high impact giving opportunities. This requires comparisons of charities that target different outcomes, such as deaths averted or gains in household consumption.

When comparing diverse charities, GiveWell makes these decisions transparent by asking staff members to provide their 'moral weights.'[34] These judgments are based on philosophical reasoning, intuition and data on beneficiary lives, and extrapolation of preferences from studies of less relevant populations. Prior to this study, there was some evidence exploring how beneficiaries trade-off between different aid interventions (for example Shapiro, 2019), or rank different policy priorities (for example Tortora, 2009). However, there was a clear lack of data on how potential beneficiaries of such interventions trade-off between different outcomes. This study represents a step to fill this gap for strategic international development decision-making.

APPROACH

IDinsight undertook this study of beneficiary preferences for GiveWell across four diverse regions in Ghana and Kenya in 2019. We surveyed respondents from low-income households (see Appendix 1 for a description of our sampling approach). In addition to 1,846 'poor' respondents, we also surveyed 246 'wealthy' respondents from wealthier households in the same communities. These 'wealthy' respondents are not included in the main estimates, but are included in certain analyses.[35]

Preferences

Our three primary methods (see Table 1) were designed to capture preferences informing two components of GiveWell's moral weights:

- **Component 1:** the value assigned to averting the death of an individual relative to doubling consumption for one person for one year.
- **Component 2:** the value assigned to averting the death of an individual under 5 relative to an individual 5 and older.

[34] Moral weights are subjective numerical values that weight the trade-off between deaths averted and gains in household consumption. Once staff members have provided their individual moral weights, the median across the group is taken to inform their cost-effectiveness model. https://www.givewell.org/how-we-work/our-criteria/cost-effectiveness/comparing-moral-weights

[35] Specifically, they are only included in the analyses of the relationship between wealth level (consumption per capita) and VSL and life satisfaction, as well as subgroup analyses for the choice experiments.

Our methods ask questions from two perspectives: individual (regarding respondents' own outcomes), and community (regarding outcomes affecting poor households within respondents' own community).

Table 6. Summary of three primary methods.

Method	Component	Perspective
Value of Statistical Life (VSL): Contingent valuation method eliciting willingness to pay for small reductions of mortality risk for self/child.	Data to inform both 1 & 2	**Individual:** Respondents are asked for preferences over outcomes for themselves and their own children. This approach was designed to align with the VSL literature.
Monetary value of life: Choice experiment measuring preferences for programs which save lives, and programs which increase household consumption, in their community.	Data to inform 2 only	**Community:** Respondents are asked to take the perspective of a decision maker and make trade-offs on allocating resources within their community. This approach was designed to mirror the trade-offs faced by a resource allocator, such as GiveWell.
Relative value of age groups: Choice experiment measuring preferences for saving lives of people of different age groups.	Data to inform 1 only	

These methods were refined over the course of 12 months of piloting, including consultation with experts in the field of VSL and preference elicitation. The three primary methods were chosen as they are directly relevant to GiveWell moral weights, and we had reasonable confidence they would produce reliable and internally valid data at scale.[36] For smaller subsamples, we varied the content or framing of the primary methods to collect data on additional areas of interest.[37]

We conducted accompanying qualitative research to capture the moral reasonings used by beneficiaries when making trade-offs related to Component 1 and 2. This work provides context to the quantitative results and illuminates how beneficiaries may be approaching these trade-offs differently to decision makers. As the three primary methods are dependent on respondent interpretation and understanding of the presented scenarios, the qualitative data also allows us to examine the credibility of quantitative responses, and identify potential survey biases.

The results of our work to capture preferences and their supporting reasonings are presented in Section 1: Value of Increasing Consumption Relative to Saving Lives, Section 2: Relative Value of Individuals Under 5 and 5 and Older, and Section 3: Moral Reasonings. These results are supplemented by additional data in Appendix 2-5 and the results of secondary methods presented in Appendix 9.

Alternate approaches

Our methods capturing preferences rely on the conversion of the benefits of a program into a unifying monetary measure.[38] This is the most prominent approach to this type of trade-off, but there are a number of concerns (both methodological and philosophical) which may mean decision makers do not

[36] See the IDinsight report, "Beneficiary Preferences: Proposal for Scale-up 2019", for more details.
[37] This includes two alternate framings of our choice experiment, one in which a comparison is made between life-saving programs and education programs and one in which the trade-off between averting deaths and increasing consumption is presented more directly.
[38] I.e. we infer a monetary value for a life saved, allowing us to compare to a monetary benefit, doubling consumption.

want to rely solely on this approach.[39] We therefore collected data relevant to alternate approaches to this problem, that may be used in conjunction with, or instead of preferences.

We captured a measure of subjective wellbeing from all our respondents, allowing us to look at correlations between factors in beneficiary lives and their life satisfaction scores. We also collected qualitative data to better understand the thought process when respondents discuss their wellbeing, and the most influential factors on their overall life satisfaction. This data could ultimately also be used to inform moral weights, by converting each of the benefits (i.e. increasing consumption, averting a death) into units of subjective wellbeing rather than money.[40]

The results of our work to capture subjective wellbeing of respondents are summarized in Section 4: Subjective Wellbeing, and supplemented by Appendix 6.

Finally, we conducted secondary data analysis of a large economic survey in Kenya, and collected primary qualitative data exploring the burden of death in beneficiary households as part of our main survey. This data can be used to create a model of the economic value of an individual to those around them. This may be purely economic, based on net present and future economic contribution to the household (similar to actuarial approaches). Or, it may expand to quantify the economic and emotional impact on a household when a family member dies. Again, this provides an alternate, non-preferences-based approach to convert the benefit of averting a death into monetary units.

The results of this work are summarized in Section 5: Economic and Emotional Impact of Death, and supplemented by Appendix 7.

[39] There are methodological concerns that the approaches used do not capture true preferences, for both revealed and stated preference studies. For example, in revealed preference, while this captures actual behaviour, it is often in scenarios where individuals are not fully aware of the risk associated with their decision. Meanwhile, in stated preferences, the framing and understanding of the question posed can create considerable bias on the results. Ethical concerns raised relate to how welfare is defined – some argue that satisfying individual preferences is less important that maximising subjective wellbeing across a population, regardless of individual preferences (see Plant, 2018).

[40] Collecting enough data to make this conversion was beyond the scope of this study, so we instead present the findings and give an overall idea in which direction it might move moral weights.

SECTION 1: VALUE OF INCREASING CONSUMPTION RELATIVE TO SAVING LIVES

TAKEAWAYS

- Using the individual perspective, **the estimated median VSL is $40,763 for individuals under 5 and $33,798 for individuals 5 and older.** This is at the upper end of values predicted by existing literature, and higher than current GiveWell moral weights (particularly for individuals under 5).
- Using the community perspective, **the implied value is $91,049 for individuals under 5 and $47,644 for individuals 5 and older.**
    - 38% of respondents chose averting the death of an individual under 5 regardless of how many cash transfers to community households were offered (up to $10 million worth of cash transfers).
    - On the other hand, among the 54% of respondents willing to trade-off between life and cash transfers, a large proportion switch to cash transfers at an implied value below $30,000. (The remaining 8% of respondents always chose the program that provided more cash transfers over the program that saved more.)
- Taken together, these findings indicate that beneficiaries appear to value averting death higher than is currently represented by the GiveWell staff moral weights.
- If these findings were to be incorporated into GiveWell's moral weights, they would increase the relative cost-effectiveness (CE) of existing life-saving top charities (e.g. AMF, Helen Keller) and make it easier for future life-saving opportunities to reach GiveWell's CE threshold.

This section details the methods and results of the study related to the value of increasing consumption relative to saving lives. It begins with the individual perspective approach, then the community perspective approach, before concluding with a comparison and discussion of implications.

INDIVIDUAL PERSPECTIVE

To capture the individual perspective on the value of averting death relative to increasing consumption, the most relevant common approach is to measure the Value of Statistical Life (VSL).

There are two approaches to estimate VSL: stated preference studies and revealed preference studies. Stated-preference studies rely on surveys to elicit respondents' willingness-to-pay (WTP) for an outcome in a hypothetical scenario. Revealed preference studies infer the value of nonmarket goods from observed behaviors in relation to market goods.[41] Due to the absence of data on the target

[41] Studies of this kind most often rely on extensive job market data, inferring VSL from increases in wages employees receive for accepting a greater risk for death in the workplace (Viscusi and Aldy, 2003).

population necessary for a revealed preference approach, we used a stated preference approach.[42] This involves the presentation of hypothetical scenarios in which WTP for mortality risk reductions is captured.

Specifically, we:

1. Introduce respondents to the scenario, in which they are told to imagine a hypothetical disease is affecting their community. Their risk of dying from this disease is 20 in 1000.
2. Introduce the vaccine/medicine (randomized) that treats this disease, and reduces this risk from 20 in 1000 to 5 in 1000 or 10 in 1000 (randomized) over the next 10 years.
3. Capture respondents WTP for this vaccine/medicine. They are told they can pay in small installments of their choosing over the 10 years of the risk reduction.

To ensure respondents understand the presented risk reductions, we used visual aids and conducted a training module on understanding small probability with every respondent. For more details on the precise set of questions used, see Appendix 2.

RESULTS

Across data collected in Ghana and Kenya, the estimated mean VSL is $40,763 for individuals under 5 years and $33,798 for individuals 5 years and older (with a standard error of $2,201, and $6,397, respectively)**.** The central estimates are relatively precise. However, this central estimate varies when alternate estimation models or samples are used (such as the different risk reduction levels, samples with varying levels of understanding, and weighting).[43]

The VSL captured in our study is substantially higher than the derived current GiveWell staff median[44] and at the upper end of the VSL extrapolations for the average beneficiary based on existing literature.[45]

[42] Revealed preference methods do not transfer well to a lower income context due to a lack of relevant datasets. Where datasets are available, they are heavily prone to selection bias as they rarely contain data on informal employment and so can miss information from the poorest households. For instance, Leon and Miguel (2017) estimated VSL in Sierra Leone by assessing travel decisions, but only captured data from a high-income sample of African travelers that are not representative of GiveWell beneficiaries. Kremer et al. (2011) studied implied VSL by examining willingness to travel to use improved water sources in rural Kenya. While the context of this study is relevant to GiveWell, it is unclear whether respondents in the study had enough information on risk levels to make an informed decision, resulting in a low estimate of VSL.

[43] For example, the full range around the central estimate for adult VSL depending on the model used is $24,440 (using the 10/1000 risk reduction, Model B) to $41795 (using the 5/1000 risk reduction, full sample). The ranges around the estimates for under 5 and over 5 VSL can be found in Appendix 2.

[44] To get the current GiveWell staff median, we convert moral weights into a derived USD. We used a conversion factor of $286 (annual consumption for a single individual per year) based on GiveDirectly's 2011 study in Kenya. For more discussion of this assumption see Appendix 8: Aggregated results.

[45] VSL among the population of the USA is $9 million (Robinson et al. 2016). Using USA nominal GDP per capita of $62,641 (2018) and elasticities of 1.0, 1.5 and 2.0, the projected VSL values of adult beneficiaries with an annual per capita consumption of $286 are $41,091, $27,765, and $1,876 respectively. Our mean of $35,733 for adults, places us at the upper end of these extrapolations.

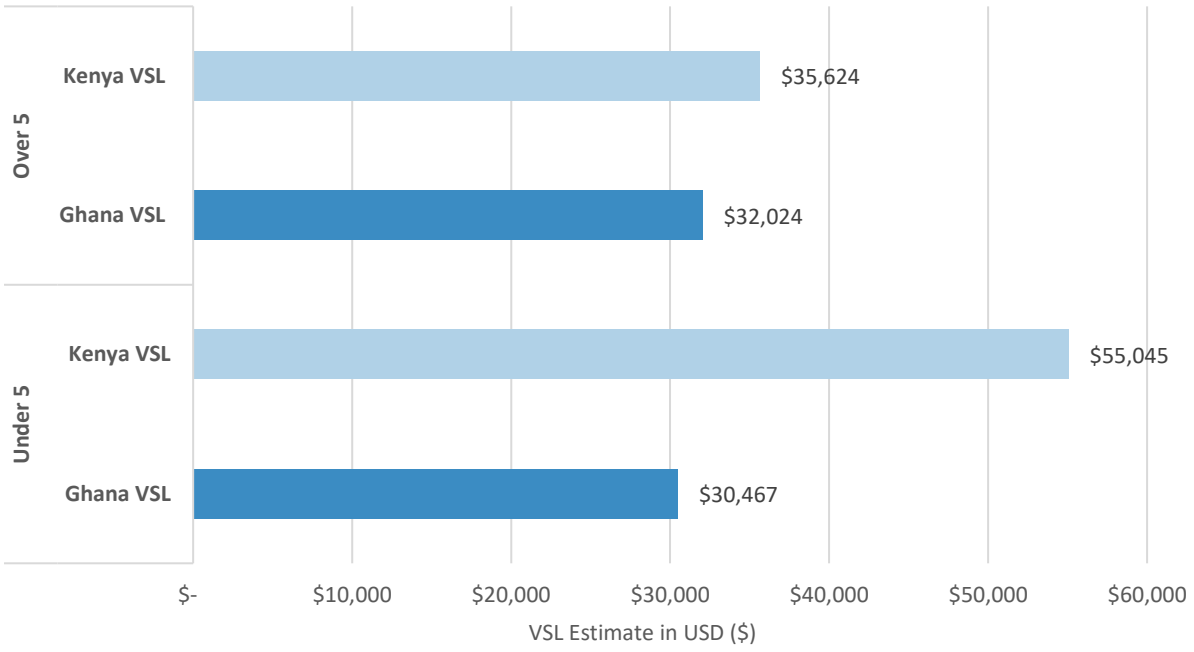
Table 7. VSL results, by age group and country, including age adjusted under 5 and 5 and older values.

	Full Sample*			Ghana			Kenya		
	N	Mean VSL (USD)	SE (USD)	N	Mean VSL (USD)	SE (USD)	N	Mean VSL (USD)	SE (USD)
I. Raw VSL Results									
Children Under 5	222	$40,763	$6,397	129	$30,467	$4,358	93	$55,045	$13,936
Children 5-18	255	$31,713	$3,331	122	$33,912	$6,189	133	$29,695	$2,944
Adults	544	$35,733	$2,909	297	$31,248	$4,216	247	$41,125	$3,900
II. Age Adjusted Results									
Individuals under 5	222	$40,763	$6,397	129	$30,467	$4,358	93	$55,045	$13,936
Individuals 5 and older	799	$33,797	$2,201	419	$32,530	$3,695	380	$35,624	$2,469

**Our full sample mean takes an average across all the observations. It is not a population-weighted average of Ghana and Kenya means.*

There is considerable variation in country-level VSL (see Figure 1)**.** In Kenya, VSL is on average higher, and there are much wider variations between under 5 and 5 and older values. Meanwhile in Ghana, VSL is lower and there is limited variation across age groups. Regional-level variation is even more substantial. This may be in part due to the relatively small sample size included once we reach this level. For individuals under 5 years, Migori (Kenya) has the highest VSL ($78,191) while Jirapa (Ghana) has the lowest VSL ($20,864). This finding highlights the fact that VSL varies widely by geography, making direct extrapolation of results from one context to another less reliable.

Figure 1. IDinsight VSL results, by country.



Across VSL data collected in Kenya and Ghana, respondents' consumption is positively correlated with their WTP – **VSL is higher for individuals with higher consumption.** We found a within sample income elasticity of 0.34. This means that for every 1% increase in consumption there is a 0.34% increase in VSL.[46] This is lower than the cross-country income elasticity predicted for LMICs of 1 to 1.5, (Robinson et al, 2019). Previous studies have also found that within country income elasticity is lower than cross-country, which could explain the results (Masterman and Viscusi, 2018).

Beyond income and regional variations in VSL, we also found that VSL was significantly correlated with respondent age (decreases with age), gender (lower for women), and household size (higher for adults living in larger households). We found no VSL correlation with literacy, religion, or life satisfaction.[47]

Our VSL results compares to the literature in the following ways:

1. We have passed the theoretical tests required to consider the results internally valid. Specifically, our aggregate data passed the weak external scope test (respondents pay statistically significantly more for higher risk reduction on average) and passed the construct validity test (wealthier respondents pay statistically significantly more for the same risk reduction). These tests are commonly used in the stated preference VSL literature to establish valid VSL results.
2. Despite sampling respondents with low education and literacy levels, understanding of small probabilities (with training) was on par with VSL literature in LMICs (most respondents understood the scenarios we presented).[48]

For more details on the validity tests of the results, see Appendix 2.

COMMUNITY PERSPECTIVE

Choice experiments have been used frequently in LMICs to capture how individuals trade-off between different services or health-states.[49] To capture the community perspective on the relative value of life to increasing consumption, we ask respondents to trade-off between saving lives and giving cash transfers within their community.[50]

[46] This comes from a regression on log VSL (in nominal USD) on log annual consumption per capita (in nominal USD), controlling for respondent characteristics (gender, age, urban location, literacy, religion, household size, self-reported long-term health conditions, life satisfaction ladder, and region dummies).

[47] These results come from a construct validity test, regressing VSL (in nominal USD) on log annual consumption per capita (in nominal USD), a dummy for risk reduction levels (5/1000 or 10/1000), and respondent characteristics as in the income elasticity regression.

[48] 58% respondents answered the four core questions on small probability and risk reduction understanding correctly the first time, and 80% got all four questions correct after at most one additional explanation on any question. Further description of how our results compare with understanding levels in the literature are found in Appendix 2.

[49] For example, the approach has been used previously to understand how individuals value different aspects of a health intervention in rural Bangladesh (Moborak et al., 2012). It has also been used extensively in LMICs by the Global Burden of Disease studies to capture comparisons of disability states (Salomon et al., 2012). Finally, similar choice experiments have also been used in the experimental philosophy literature to capture moral trade-offs on sensitive topics in the US (Elias et al., 2016).

[50] We piloted the community perspective method over a period of 1.5 years leading up to the scale up exercise and refined our approach to address the limitations we observed. Most notably, we changed the framing such that programs in both choice sets save lives and provide cash transfers at different levels. This is because respondents demonstrated sensitivity to framing when asked to directly compare the program that saves lives with another that provides cash transfers.

We conducted a choice experiment in which respondents' trade-off between two hypothetical interventions:

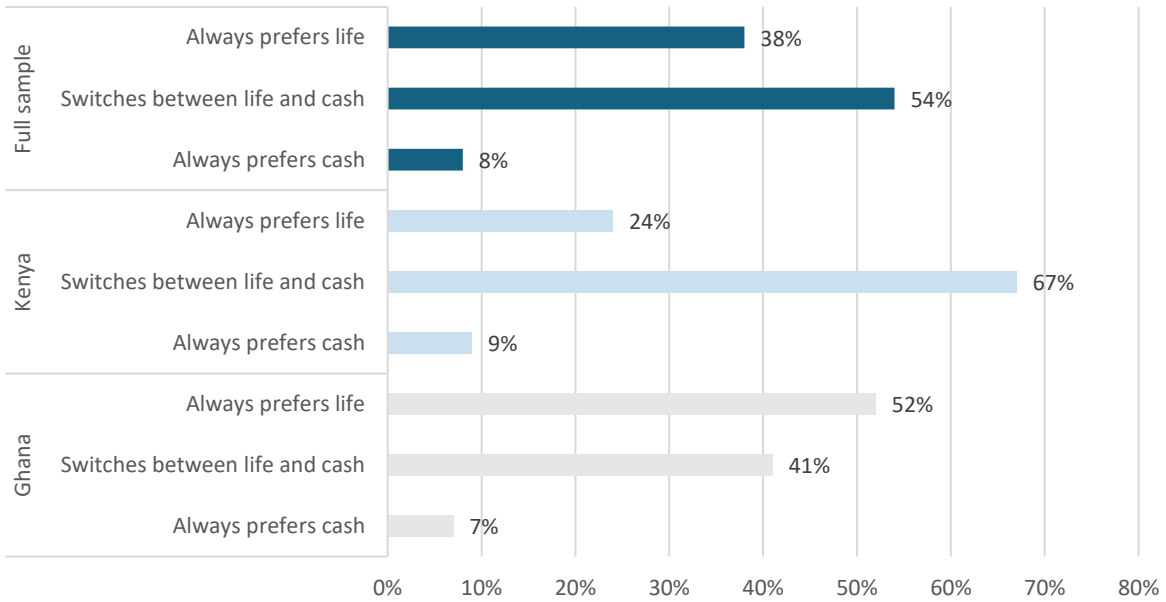
> *"Program A saves the lives of 6 children aged 0-5 years AND gives $1,000 cash transfers to 5 families. Program B saves the lives of 5 children aged 0-5 years AND gives $1,000 cash transfers to [X] families. Which one would you choose?"*

We varied the value of *X*, both within and across respondents to capture the switching point between cash transfers and saving the life of an extra child under 5 across the population. As respondents make choices, the presented programs become more extreme to explore whether switching occurs when the difference is starker.[51] The full method specification can be found in Appendix 3.

RESULTS

Across our sample of low-income respondents in Ghana and Kenya, we found that 38% of respondents always chose the program that saved more children's lives over any number of cash transfers offered (up to the presented maximum of $10 million).[52] A further 8% of respondents always chose the program that provided more cash transfers over the program that saved more children's lives. The remaining 54% of respondents switched from the program that saved more lives when the number of cash transfers given by the alternate program was progressively increased.

Figure 2. Distribution of preferences between those that always prefer a program that provides more cash transfers, those that always prefer programs that save lives, and those that switch as the level of cash transfers vs lives saved varies.



[51] For example, if respondents always choose the program that saves lives, the number of cash transfers given by the alternate program is progressively increased.

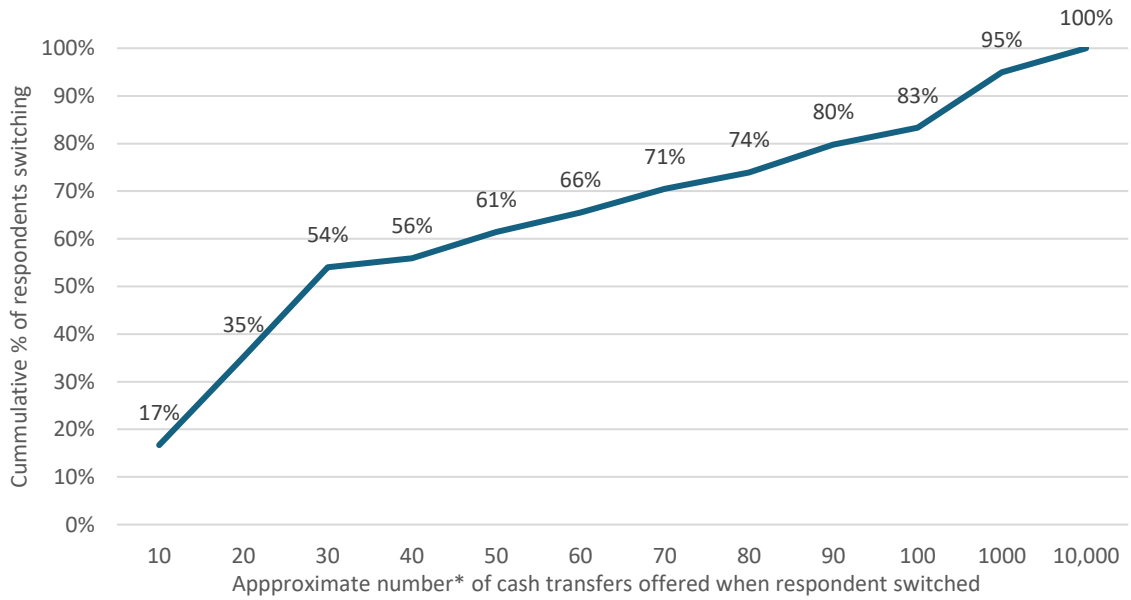
[52] We chose to use $10 million as the highest value in our questions as it is the high end of US VSL estimates.

There were substantial country-level differences in the distribution of switching points**.** Most notably, **52% in Ghana always prefer life-saving programs, no matter how much cash is offered, compared to 24% in Kenya** (Figure 2).

The central estimate[53] of the monetary value of life from the community perspective method reveals a ratio of 91.05 (SE: 44.19) between cash transfers and saving lives of individuals under 5 years. Across the full sample, **we find that saving the life of a single child under 5 in the respondents' communities is equivalent to giving cash transfers of $1,000 to ~91 poor households, or an implied value of $91,049 for a child under 5 years old** (SE: $44,188). The implied value of an individual under 5 ($91,049) is substantially higher than the derived current GiveWell staff median ($13,442).

This result is driven by the proportion of respondents who do not switch regardless of the number of cash transfers offered.[54] Among individuals who do switch, many respondents prefer the program offering cash transfers at a relatively low number of transfers. Approximately half of the 'switchers' in our sample do so at an implied value of $30,000 or less (Figure 3).[55]

Figure 3. Cumulative % of respondents switching to the program offering more cash transfers, by the number of cash transfers offered (n=905 switchers, who passed dominance and consistency check).



**We can only provide an approximate number due to the overlapping nature of the choice sets. For example, one respondent might be offered 20, 40, and 60, while another is offered 30, 50, and 70. To form this graph, we have assumed for all that the switching point is at the upper bound of the choices offered. Graphs portraying the raw switching intervals are included in Appendix 3.*

[53] To get the central estimate of the monetary value of life from the varied switching and non-switching patterns, we combined the results from each respondent's first three choices; we did not include choices from latter questions in the estimation (but only used them to illustrate the switching points) as these choices depend on earlier choices and hence are endogenous, leading to bias if included in the estimation.

[54] For them, the value of saving the life of a single child under 5 in the respondents' community is between $10 million and infinity.

[55] It is not possible to extract a precise median from our results, due to the overlap in choice sets presented to individuals.

The importance of the non-switchers on our estimation is reflected in the stark differences between country-level estimates. In Ghana, where 52% of respondents always prefer life-saving programs the central estimate is $200,877 (SE: $1,352,399). Meanwhile for Kenya, where 24% of respondents always prefer life-saving programs, the central estimate is $14,499 (SE: $67,173).

For most respondents non-switching seems to reflect a true strong preference for life-saving interventions. As explored further in Section 4: Moral Reasoning, respondents with strong preferences for life-saving interventions gave a wide range of justifications. Additionally, only a small number of respondents expressed any negative perception of cash, which could bias them against cash transfers relative to another consumption increasing intervention.[56]

Additional analysis, presented in Appendix 3, showed that non-switchers, who always pick life-saving interventions, are demographically different from switchers. They are more likely to be illiterate, come from a large household with more children, and there may be some correlation with religious views. We also found that non-switchers place higher relative value on children under 5, compared to other age groups.

We found good evidence that respondents understood this method. Nearly all respondents (95%) passed a basic understanding test, 92% demonstrated consistent preferences across their choice-set, and 88% passed both types of tests (see Appendix 3, for more details).

DISCUSSION OF RESULTS

In this subsection we discuss key results from the above analysis, namely: between country differences, between approach differences, aggregating results across approaches, and comparison of the study results to the literature and GiveWell's current moral weights.

Between country differences: comparing results from Kenya and Ghana

Across both the individual and community perspective, we found substantial geographic variance in preferences captured. Due to overall confidence that our approaches were implemented consistently in both contexts, we believe this reflects a true difference in preferences.

Subjectively, one of the biggest differences we noted between the communities we worked in in Kenya compared to those in Ghana, was the level of within community inequality. In Kenya, there were clear differences between the wealthiest and the poorest in each community, and we rarely struggled to identify 'wealthy' households to interview. However, the opposite was true in Ghana – here we found very little variation in within community wealth. We struggled to find *any* wealthier respondents in rural areas, and depended on urban areas to supplement this sample. Our hypothesis is that the experience with inequality *may* influence how people view the relative importance of improving living standards. However, this is currently based only on our own subjective observations and warrants further research to better understand geographic and cultural differences in preferences.

[56] 6.2% of respondents stated that poor households would misuse cash transfers. We asked respondent what poor households in their community would do with a $1,000 cash transfer. We only counted responses as "misuse" if they mentioned squandering, wasting, gambling, or drinking; we did not count other answers like entertainment or marrying another wife as it is less clear whether these are perceived negatively.

Between approach differences: comparing results from the individual and community perspectives

Across both countries, the community perspective central estimate of the monetary value of life is higher than the individual perspective (Table 8).

Table 8. Technical comparison of Individual and community perspective results on the absolute value of life.

	Individual Perspective	Community Perspective
Value USD – Under 5	$40,763	$91,049
Value USD – 5 and older	$33,798	$47,645
Precision of estimate	Small SEs –relatively high precision in the estimate across the full sample.	Low precision, high SEs.
Sensitivity to estimation approach	Reasonable sensitivity to the risk reduction level that is used to calculate VSL, and the selected model.	Relatively high sensitivity to chosen estimation method, due to the impact of 'non-switchers' on the model.

The difference in results likely reflects important differences in the design of each method. The key theoretical differences between the methods are:

1. The individual perspective asks respondents to value their life or their child's life, whereas the community perspective asks respondents to value community members. (On this dimension, the community perspective may be more relevant to moral weights as it asks beneficiaries to make a similar trade off to that made by GiveWell staff.)
2. The individual perspective involves reducing mortality risk by a small amount, whereas the community perspective involves saving lives with certainty. (On this dimension, the individual perspective may be more relevant to moral weights as GiveWell top charities reduce the risk of mortality.[57])
3. The individual perspective is based on WTP and, even with integrated steps to overcome a likely liquidity constraint,[58] respondents are limited in what they are willing to pay by their income (i.e. what they are able to pay). In contrast, in the community perspective, respondents make trade-offs as decision-makers, and so their own income should not constrain their decisions.

Issues that may arise in implementation of the two approaches that could lead to further differences include:

[57] We considered using a version of the community perspective involving reducing mortality risk but decided against it given the added complexity of this formulation.
[58] 'Liquidity constraint' refers to a constraint on an individual's ability to pay for something, due an inability to borrow caused by imperfect capital markets. To overcome this constraint, we ask respondents for their WTP in small instalments over 10 years such that each individual hypothetical payment is much less influenced by any liquidity issues.

4. Potential greater social desirability bias of the community perspective, leading respondents to be more likely to choose life-saving interventions (as it explicitly asks respondents to choose between saving lives and other interventions).
   - We see very little evidence of this in our qualitative work, but it is generally difficult to measure social desirability bias.
5. Potential impact of negative perceptions of cash transfers in the community perspective leading respondents to systematically undervalue of cash.
   - We see some evidence of respondents using this justification to choose life-saving programs (see Section 3: Moral Reasonings). However, in our quantitative survey, only 6% of respondents expressed the view that cash transfers would be misused, so we do not believe this is unduly biasing our results.

Aggregating results across approaches

Note: Developing a single estimate of moral weights estimate from the results of this study is not straightforward and requires a number of significant assumptions. Here we present our best central estimate of preferences. Readers may want to adjust these assumptions, or take a different approach altogether.

In order to form a single central estimate of beneficiary preferences that can be incorporated into GiveWell's CEA model, we aggregate results across our main approaches by weighting the individual and community perspectives equally, then convert our value from USD into moral weights (for results, see Table 9). The full process for aggregating and converting results is outlined in Appendix 8.

Our aggregated values of averting a death are $40,721 for individuals 5 and older, and $65,906 for individuals under 5. Fully incorporating[59] these values into moral weights would:

- **Increase by a factor of 1.7 (for individuals 5 and older) and by a factor of 4.9 (for individuals under 5)** the value placed on averting death relative to doubling consumption.
- **Increase the value** placed on individuals under 5 relative to individuals 5 and older **from 0.6 to 1.7.**

This would lead to a higher relative cost-effectiveness of charities whose good is achieved primarily by averting the death of young children (e.g. Helen Keller International, Malaria Consortium, and Against Malaria Foundation see Figure 17, Appendix 8). Charities whose good is achieved primarily by increasing consumption (e.g. Sightsavers and GiveDirectly) would become comparatively less cost-effective. The change in moral weights would also make it substantially easier for other charities whose main outcome is saving lives, particularly of children under 5, to reach and surpass GiveWell's cost-effectiveness threshold (2-3x as cost-effective as cash).

[59] By fully incorporating we specifically refer to giving 100% weight to our aggregated results, such that moral weights solely reflect the central estimate of beneficiary preferences.

Table 9. Aggregated beneficiary preferences, in comparison to the current median of GiveWell staff moral weights.

	Derived value (USD)		Moral Weights	
Value assigned to…	Aggregated beneficiary results	Current GiveWell median	Aggregated beneficiary results	Current GiveWell median
Averting the death of an individual under 5	$65,906	$13,505	230	47
Averting the death of an individual 5 and older	$40,721	$24,606	142	85
Doubling consumption for one person for one year			1	1

Comparison of the study results to the literature and GiveWell's current moral weights

Respondents placed substantially higher weight on averting death relative to doubling consumption than predicted by most of the standard methods of extrapolation from HIC data, and than GiveWell's current aggregate moral weights (Figure 4 and Figure 5).

Figure 4. Implied Value USD for averting the death of individuals 5 and older. Comparison of study results, literature priors, and GiveWell 2018 moral weights.

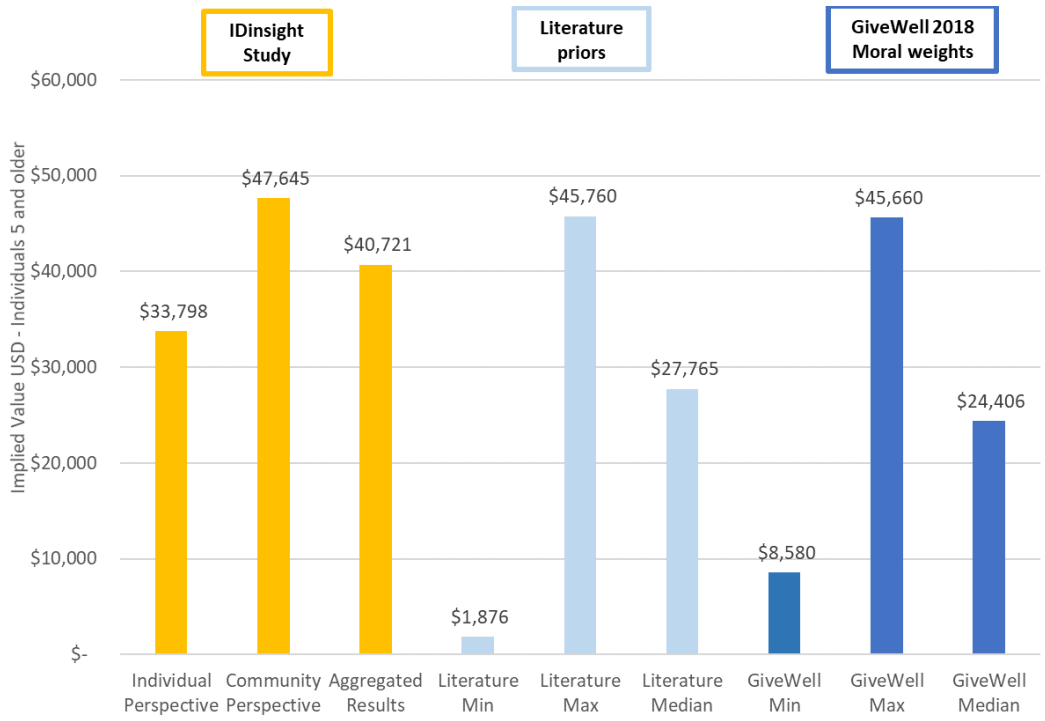
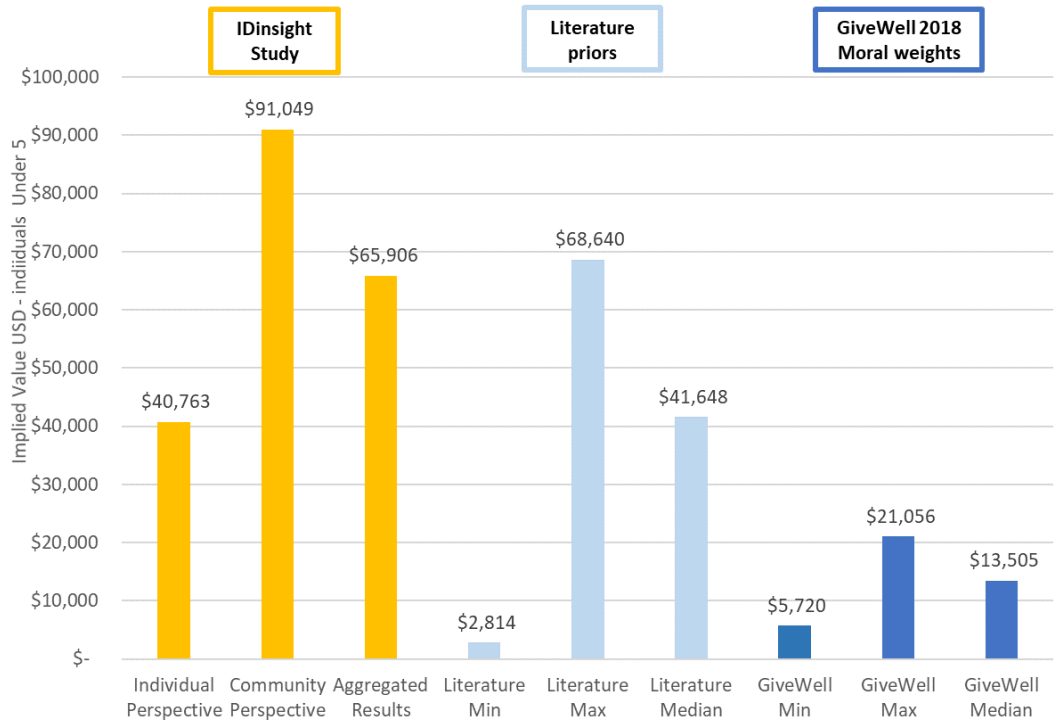


Figure 5. Implied Value USD for averting the death of individuals under 5. Comparison of study results, literature priors, and GiveWell 2018 moral weights.



**As there is less pre-existing literature available providing a precise valuation for children under 5, we take the min, max, and median values for adults, and apply the Robinson et al. (2019) central estimate for the relative value of children compared to adults (1.5x).*

SECTION 2: RELATIVE VALUES OF INDIVIDUALS UNDER 5 VS 5 AND OLDER

TAKEAWAYS

- We find an estimated value of individuals under 5, relative to individuals 5 and older of:
  - 1.2 when taking the individual perspective,
  - 1.9 (excluding individuals above 40 years) and 3.7 (including individuals above 40) when taking the community perspective.
- These ratios are consistent with the range predicted by the literature for this population, but in contrast to GiveWell's current staff aggregate ratio (0.55), which places more value on individuals 5 and older.
- If study results on the relative value of lives were to be incorporated into GiveWell's moral weights, it would increase the cost-effectiveness of the charities that primarily save the lives of children under 5 (e.g. AMF, HKI). Additionally, it would make it easier for other charities that primarily target children under 5 (and that aren't top charities yet) easier to reach this cost effectiveness threshold.

This section describes beneficiaries' relative valuations of children who are under 5 years, and individuals who are 5 years and older. This trade-off is central to GiveWell's moral weights framework as charities often serve individuals of different ages. A number of their charities specifically target or benefit individuals under 5, such as AMF (the benefits of bed nets are greatest in this age group) and Malaria Consortium (who provide seasonal chemoprevention for Malaria specifically to children under 5). For this reason, in their current cost-effectiveness model most distinctions are made between the under 5 and 5 and older age groups. In this report we adapt our estimates to align with these two age groups, but our data does provide more granular detail on the relative value of ages which may be of future interest.

We estimate the value of each age group using two methods: (a) individual Value of Statistical Life (VSL) and (b) a community perspective choice experiment. Both approaches have been used in the literature previously. In the VSL literature, the typical approach (which we have followed here), is to ask for a parent's WTP to avert mortality risk for their own child. This has been completed in a number of HIC studies, as summarized by Robinson et al. (2019), but to the best of our knowledge has not previously been conducted in any country in Africa. Estimates from HICs typically find that children are valued over 1.5 times adults, but the range around this estimate is large (from 0.6 to 2.9).

The community perspective choice experiment follows an approach designed by Johansson-Stenman in recognition of concerns with using the VSL approach to capture the relative value of different age groups. Of note, it has been conducted previously in on LMIC (Bangladesh: Johansson-Stenman et al. 2009), where the implied ratios were found to be similar to those captured via the VSL approach.

INDIVIDUAL PERSPECTIVE

Similar to Section 1, we use the VSL method to capture the relative value of individuals under 5 to individuals 5 and older, from the individual perspective. Following the most common approach in the literature, we capture the WTP for a vaccine for oneself, and compare that to the WTP for the same vaccine for one's child. For child VSL, we randomly selected one of the children in the household for whom the respondent was the biological parent or main caregiver. We weighted this selection, such that ~50% of the time the child selected was under 5 and the rest the child was 5 and older. We randomized the order in which we asked questions about the respondent themselves, and their child, to control for any ordering effect. Here we present VSL for each group: children under 5, children 5 and older but under 18, and adults. We combine the estimates for children 5 and older and adults according to population distribution to give a single estimate comparing under 5 and 5 and older.

RESULTS

Across data collected in Ghana and Kenya, we estimate a ratio of 1.2 between the value of individuals under 5 and 5 and older, see Table 10**.** This means that individuals under 5 are valued 20% higher than individuals 5 and older on average. This ratio is within the realm of ranges predicted by the HIC literature (0.6-2.9), and very similar to the GiveWell conventional column ratio (1.23). However, the median ratio of staff member moral weights for under 5 is 0.55. In contrast to our results, GiveWell staff members have previously placed greater weight on individuals 5 and older.

Table 10. VSL Results by age and country, and resulting adjusted Under 5 vs 5 and older ratios.

	Full Sample			Ghana			Kenya		
	N	Mean VSL (USD)	SE (USD)	N	Mean VSL (USD)	SE	N	Mean VSL (USD)	SE
I. Age Adjusted Results									
Individuals under 5	222	$ 40,763	$ 6,397	129	$ 30,467	$ 4,358	93	$ 55,045	$ 13,936
Individuals 5 and older	799	$ 33,797	$ 2201	419	$32,530	$ 3,695	380	$35,623	$ 2,469
II. Implied Relative Ratio (Under 5:5 and older)									
Ratio	1.21			0.94			1.55		

There are substantial country-level variations in this finding (see Table 10). In Ghana, VSL is relatively consistent across age groups, resulting in a ratio close to 1 (0.94) in which slightly more value is placed on individuals 5 and older. In contrast, across our sample in Kenya, children under 5 are valued substantially higher than children 5 and older ($55,045 vs $29,695). Of note in Kenya, however, is the relatively small number of observations for children under 5 that were included in the final analytical model, resulting in a less precise estimate.[60] In the full sample, we take a the mean across all

[60] By chance, a disproportionately large number of observations for individuals under 5 were excluded due to poor understanding of the respondent.

observations, so slightly more weight is therefore given to the Ghana results as they contribute more observations in the final model.

As noted in Section 1, we are overall confident in our VSL results as we passed the theoretical tests required to consider the results internally valid (see Appendix 2). However, from our qualitative research, we identified a common misunderstanding that may have affected the derived ratio of child vs adult VSL. A number of respondents noted giving lower values for children under 5, because they expect children's medicines to be cheaper than adult's, when purchased within their local markets. While this reasoning was not widespread across the full sample, it *was* common among individuals who were WTP less for their child's vaccine than their own. We found that approximately 37% of respondents gave this justification for paying less for a child's vaccine.[61] A possible implication of this result is that our ratio of individuals under 5 to 5 and older (1.2) is slightly lower than the true values held by respondents.

COMMUNITY PERSPECTIVE

In order to capture a community perspective on the relative value of individuals under 5 and individuals 5 and older, we ask respondents to trade-off between community-level programs that save different numbers of lives in different age groups. We conducted a choice experiment in which respondents' trade-off between two hypothetical interventions:

> *"Program A saves [100/200/300/400/500] lives of people aged [under 5/5-18/19-40/over 40], Program B saves [100/200/300/400/500] lives of people aged [under 5/5-18/19-40/over 40]. Which one would you choose?"*

We varied the attributes highlighted above (ages and quantities) both within and across respondents to capture the relative preferences of lives of different ages. This method follows very closely the protocol used by Johansson-Stenman et al. (2008, 2009) when capturing the relative value of lives of different ages in both Sweden and Bangladesh. The full method specification can be found in Appendix 4.

RESULTS

Across our full sample in Ghana and Kenya, we found that **individuals under 5 are consistently valued higher than individuals of all other age groups.** We obtained the following ranking of the age groups we presented: **under 5 years > 5-18 years > 19-40 years > over 40 years**.

[61] We asked this question to a subsample of 675 respondents, predominantly in Ghana. Of those, 197 paid less for the child vaccine, 72 of them gave the market value of children's vaccine as the main reason for this.

Table 11. Raw estimation results, and implied relative ratio of under 5 to 5 and older, from our community perspective choice experiment.

	Full sample (n=1493)	Kenya (n=703)	Ghana (n=790)
I. Raw estimation results.			
Ratio of (5 to 18) to (under 5), (SE)	0.792 (0.002)	0.898 (0.003)	0.636 (0.008)
Ratio of (19 to 40) to (under 5), (SE)	0.113 (0.002)	0.200 (0.002)	-0.027 (0.009)
Ratio of (over 40) to (under 5), (SE)	-0.711(0.012)	-0.267 (0.006)	-1.523 (0.131)
II. Implied relative ratio (under 5 to 5 and older)			
Ratio of (under 5) to (5 and older) excl. over 40	1.9	1.6	2.7
Ratio of (under 5) to (5 and older) incl. over 40	3.7	2.2	-80.2

Across respondents in both countries, the calculated value of individuals under 5 years relative to individuals 5 years and older is 3.7. While this ratio captures the overall strong preference for individuals below 5 years relative to other age groups, it is highly skewed by an estimated *negative* valuation of individuals above 40 years. In both Kenya and Ghana, the ratio of individuals over 40 years relative to individuals below 5 years has a negative coefficient. Qualitative data suggests that at least some respondents consider the elderly to contribute a *net drain* on household resources,[62] although it is not clear to what extent this explains the negative valuation. Plus, it is contradicted by one of our secondary methods which did find a net positive valuation of individuals over 40 (albeit from a much smaller sample, see Appendix 9 for more details).

While this finding is comparable to the results from a similar study in Bangladesh (Johansson-Stenman et al. 2009), it has a substantial effect on the overall results of our study. Plus, we know that few current GiveWell top charities support individuals over 40, so the relative value of this group is of less relevance to mora weights. Finally, we are not confident that the negative estimates for those above 40 reflect true preferences based on the mechanics of the analytical model. When comparing any two age groups against each other, over 40 is the group least frequently selected.[63] Since choices involving those over 40 are more one-sided, the model may not have enough data to accurately estimate the value placed on those over 40, despite the apparently small standard error.

As such, **we suggest placing more weight on the ratio obtained when individuals over 40 years are excluded (1.9).** This ratio is consistent with the range predicted by the literature for this population.

Overall, we found good evidence that respondents understood this method and made informed trade-offs. Nearly all respondents (97%) passed a basic understanding test, 85% demonstrated consistent preferences across their choice-set, and 83% passed both types of tests (see Appendix 4) Beneficiaries

[62] "Their burden is big; they take more than they provide […] But my father could still farm when he was old for a while; then he fell sick for more than 2 years; and he wasn't helping the farm at all; so it was better he died than he was living." (Male, 52, Jirapa, Ghana)

[63] They were chosen 20% of the time over those under 5, 15.6% of the time over those 5 to 18, and 22.6% of the time over those 19 to 40; all other groups were chosen at least 35.7% of the time when compared with any other group.

backed-up their choices with a wide range of reasonings, and the majority demonstrated they weighed up all aspects of the choice when making their decision (see Section 4, and Appendix 4).

DISCUSSION OF RESULTS

Our two approaches yield two relatively similar central estimates of the relative value of individuals under 5 to those 5 and older: 1.2 in the individual perspective; 1.9 in the community perspective (see Figure 6). This is further supported by one of our secondary methods that captured a relative value of 1.3 (see Appendix 9). A comparison of the credibility of the results of our two approaches is found below in Table 12.

Table 12. Technical comparison of individual and community perspective results informing the relative value of individuals under 5 vs 5 and older.

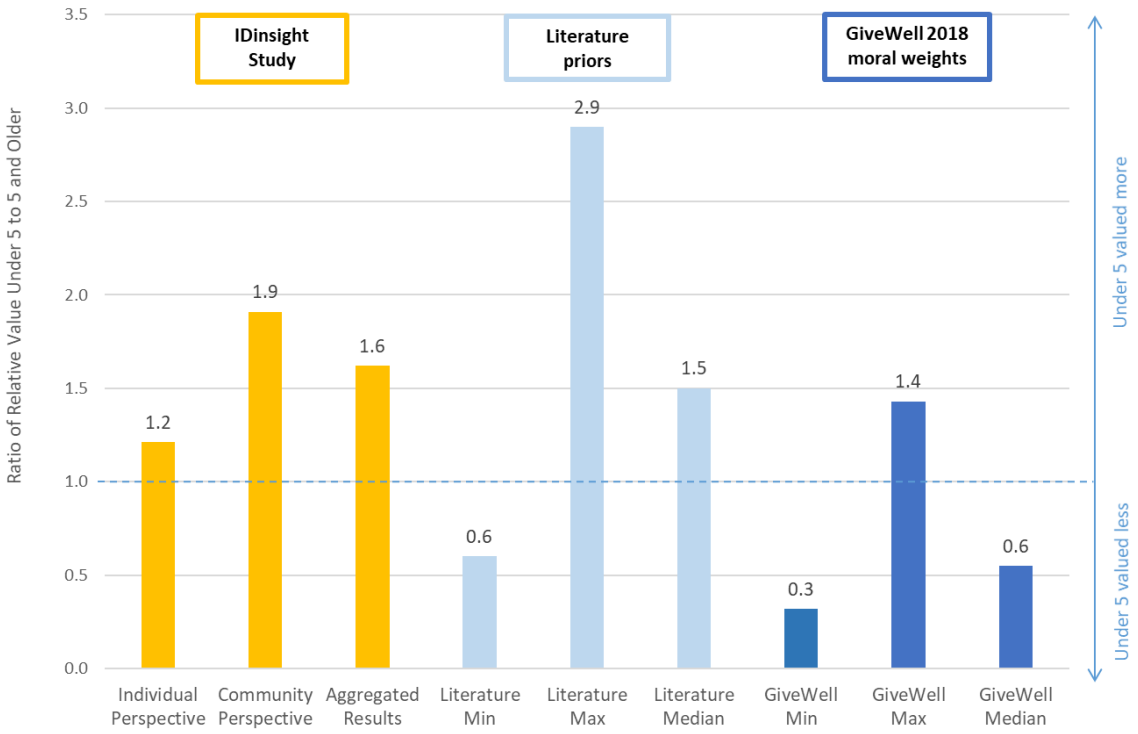
	Individual	Community
Ratio Under 5: 5 and older	1.21	1.91
Precision of estimate	Precision of the ratio is limited by the child under 5 value, which is less precise due to low final sample size in Kenya.	Precision of estimate is high.
Sensitivity to estimation approach	Low sensitivity; Very limited variation with different estimation approaches.	Highly sensitivity to the inclusion or exclusion of the valuation of individuals over 40.
Other credibility concerns	Anchoring on the market value for children's medication, may result in bias.	High negative valuation of individuals over 40 is likely not credible.

Despite the variations within and between the individual and community perspectives, **ratios from both perspectives are consistent with placing a higher value on individuals under 5**. This is consistent with literature priors about the relative value of different age groups. However, this stands in contrast to GiveWell's current staff aggregate (0.55), which places more value on individuals 5 and older (see Figure 6).

The contrast likely originates from a difference in perspective on valuing lives of different ages. A number of GiveWell staff members follow the personhood approach, placing lower value on young children, as they have not yet reached their full 'life potential.' Meanwhile, respondents tend to place high value on the potential for young children to achieve an extremely high pay-back over the course of their lives.[64]

[64] See Moral Reasonings in Section 4 below for more qualitative reasons respondents gave for placing high value on individuals under 5 years relative to other age groups.

Figure 6. Comparison of the relative value of individuals under 5 vs 5 and older, between our study, literature priors, and GiveWell's 2018 staff median moral weights.



If this finding were to be incorporated into GiveWell's moral weights, it would increase the cost-effectiveness of charities who primarily outcome is saving the lives of children under 5 (e.g. AMF, HKI). Additionally, it would make it easier for charities which save the lives of children under 5 to reach the GiveWell cost-effectiveness threshold.

SECTION 3: MORAL REASONINGS

TAKEAWAYS

- Respondents use a wide range of ethical frameworks when making moral trade-offs. These frameworks have similarities with the major moral philosophy traditions – utilitarian, deontological, and virtue ethics, but no one philosophy dominates.
- The high value respondents place on averting deaths, particularly those of young children, relative to increasing consumption, is driven by:
    - A large proportion of the sample who place high weight on the idea that life holds inherent value and therefore no amount of money is sufficient to forego the chance to save a life.
    - Respondents reasoning that there is a need to protect the potential held by all individuals to achieve high economic and social value over their life-course, particularly young children.
- Many respondents' moral reasoning differs from the approach used in GiveWell's cost-effectiveness analysis, which assumes outcomes can be compared – to many respondents it is not possible to compare the value of saving a life to other outcomes (such as increasing consumption).

OVERVIEW

To better understand how beneficiaries make moral trade-offs, we collected in-depth qualitative data on the reasons behind respondents' valuations of different outcomes and interventions.[65] We use this data to answer three questions that are key to understanding and using our quantitative results:

1. *Do beneficiaries make informed trade-offs?* Here we look at the range of factors considered by beneficiaries when responding to our questions, and whether they demonstrate a considered weighing-up of the presented options.
2. *Why do beneficiaries make certain trade-offs?* Here we look for patterns in the justifications given by beneficiaries that explain the quantitative values presented in Section 1 and Section 2.
3. *Are the decision frameworks used by beneficiaries different from those used by other decision makers?* Here we examine the high-level recurring arguments made by beneficiaries, and ask whether these are different or similar to those used in GiveWell's cost-effectiveness analysis (which relies on comparing across outcomes).

Our findings present data from 47 individual qualitative interviews, 8 focus groups, and around 1000 quantitative survey respondents. We followed a three-stage coding operation: in stage one, we lifted

[65] We captured respondents' moral justifications through various preference elicitation questions throughout the qualitative interviews. For instance, we asked respondents to pick between programs that save lives of individuals of different ages and programs that increase consumption for poor households. We then captured the in-depth thought processes respondents' use when making these tradeoffs.

core ideas from respondent data; in stage two, generated reproducible codes based on recurrent ideas; and in stage three, categorized similar codes into broader moral frameworks.[66]

We present respondent data through two heat maps (one per component). These are provided in Appendix 5: Moral Reasoning Data. Each map collates (a) the most common justifications for decisions made, (b) key quotes to help illustrate those justifications, and (c) the broader moral frameworks within which we locate these justifications. A color scale indicates the overall prevalence of each justification across the data.

Importantly, **we find that a respondent will commonly consider multiple justifications and moral frameworks in making a single moral trade-off.** Though we did not observe a common progression of justifications across respondents, we found that different question framings lead to varied responses.[67] As such, a respondent's determination is dependent on the trade-off presented, and is the product of weighing multiple moral factors (more below). Further, 'prevalence' does not refer to the number of individuals who subscribe exclusively to a moral reason, but rather to the total number of *instances* of that reason across responses.

FINDINGS

Respondents attribute their answers to our individual and community perspective methods to a range of moral justifications (see Table 35 and Table 36 in Appendix 5). A few justifications feature prominently across the sample, which we present (below).

COMMON JUSTIFICATIONS FAVOURING LIFE OVER CASH

1. Considering future social, emotional, and economic value of lives to be much *higher* than any present value of cash today

A common justification for prioritizing life (over cash) and children under 5 (versus 5 and older) is the belief that individuals hold the potential to achieve very high economic and social value over their life-course. For example, a child who is saved today could become a great leader in the future, and create more wealth than would be possible through cash transfers. One respondent describes,

> *"Children are very important—they are the leaders and economic forces of the future. The one child you save could be the one God blesses to provide and take care of the family in future." -- (Female, 26, Migori, Kenya)*

This rationale involves judgement on uncertain outcomes, as it is impossible to know whether anyone will achieve high economic and social value. Nonetheless, respondents may feel the chance of this happening is greater than the chance of cash transfers significantly improving outcomes -- as such, they prefer to bet on life than on cash.

[66] Given how subjective this type of analysis is, we got a second independent analyst to extract and analyse the same data and found that the same themes arise.

[67] We find that when the tradeoffs are more direct (e.g. imminent death versus increasing consumption), most respondents pick saving the life (consistent with the findings of our secondary methods in Appendix 9). However, when the same respondents are asked to take the position of a decision maker in the community and split resources between risk-reducing programs and increasing consumption, the same respondents are willing to pick increasing consumption.

Individuals also choose life because, in their view, satisfaction is more often derived from physical health than from monetary wealth in the long-run. For example,

> *"When you are all healthy, even if you are hungry, it is better this way. You will continue surviving as you would have survived anyway without the needs...when people are alive and healthy, everything else is better. But when people die, a gap that is impossible to fill (cannot be filled by anything else even cash) is left there."*
> *-- (Female, 42, Migori, Kenya)*

Respondents go on to say that families are more productive when they are satisfied with life, and less productive when they are stressed about illness and death. While cash can always be sought after, the emotional burden of death can stay with a family for long time.

2. Considering the lives of all individuals to be inherently priceless

Many respondents attribute a high valuation of life to an ethical rule that saving life is morally right, and equating life to a cash value is impossible and morally incorrect.

> *"When you have life, God has been faithful, and you can expect life will grow to assist the community. But if one has all these cash and doesn't have any extra life or child to spend it, then what is the profit in having all this cash?" (Female, Karaga, Ghana)*

> *"Life is very important and cannot be valued using cash" (Female, Migori, Kenya)*

There is notable religious influence in these justifications, including mentions of the "sanctity of life" and the "religious duty" to care for life.

3. Considering negative or unintended effects from cash transfers

A final group highly values saving life, due to an expressed scepticism around the effectiveness of cash interventions. These respondents note concerns that recipients would either intentionally misuse funds for non-productive goods (e.g. alcohol), or recipients would not know how to use the cash lucratively (e.g. create a venture). One respondent explains:

> *"Most poor families will misuse cash transfers if given directly-- they will use it to buy things they did not have. They will start business they do not have any experience in. They will buy the things they do not have but these will not help them." (Male, 26, Migori, Kenya)*

A number of respondents also fear that cash transfers could exacerbate community tensions if the funds are not distributed equally to everyone. Altogether, these respondents approximate to less than 10% of the sample, so we do not anticipate this view to have systematically biased the study's valuation of cash.

COMMON JUSTIFICATIONS FAVOURING CASH OVER LIFE

4. Considering the present value of cash to be *greater* than children's future potential value

A number of justifications are used for choosing cash over life and adults over children, and are mentioned by more than a quarter of our respondents. The first rationale is that a cash transfer is likely to significantly reduce poverty, in part because families can use cash in diverse ways, depending on their present needs (e.g. school fees, health expenditures, start ventures, improve house conditions). Here, the expected value of good outcomes from cash is *greater* than the expected value that a saved child will create wealth for the family in the future.

Many also mention that cash can also be used to benefit many in the household, in contrast to a health intervention that saves just one life. In absolute terms, cash transfers are more far-reaching in who they benefit in the present. Finally, some respondents note that the family has already invested cash in individuals 5 and older, and so it would be a waste to risk their lives for children under 5.

DISCUSSION

In the community perspective trade-off between life-saving and income increasing programs for community, 8% of respondents always chose cash over life, 38% always chose life over cash, and 54% switched from life to cash as the value of the cash transfer increased (see Section 1). Understanding beneficiaries' moral frameworks helps contextualize these results and indicate why beneficiary preferences may differ from GiveWell staff moral weights.

The high proportion of respondents who place extremely high value on life are informed by a mixture of future-oriented valuation of life (justification 1, above), beliefs on the inherent value of life (justification 2), and to a lesser extent, beliefs that may undervalue the benefits of cash (justification 3). These frameworks fall outside the implied approach of the GiveWell cost-effectiveness method which compares across outcomes, which explains the differences between results from these respondents and GiveWell moral weights. Meanwhile, a present-oriented framework (justification 4) describes why respondents switch from life to cash, even at somewhat low numbers of cash transfers. This framework is more aligned to the GiveWell cost-effectiveness method and results in value of life estimates similar to GiveWell moral weights.

The moral reasonings may also help describe some country-level differences in the community perspective. Significantly fewer respondents always chose life in Kenya compared to Ghana (24% v 52%), and there are more switchers in Kenya (67% v 41%). Consistent with this breakdown, our qualitative analysis showed significantly fewer incidences of a future-oriented framework (justification 1) in Kenya, and more instances of a present-oriented framework (justification 3).

SECTION 4: SUBJECTIVE WELLBEING

TAKEAWAYS

- Across poor respondents in Kenya and Ghana, the average life satisfaction ladder score is 2.8 (where 0 is the lowest and 10 is the highest score).
- Respondents with higher consumption have higher life satisfaction ladder scores; doubling consumption is associated with being 0.4 steps higher on the ladder.
- When describing different points on the ladder respondents most often referred to levels of money and material goods. In contrast, health states were mentioned much less often with regards to life satisfaction. Having a health condition was associated with being 0.3 steps lower on the ladder.
- Overall, taken alone, these findings suggest that consumption is of greater relative importance to wellbeing of respondents than their preferences (described in Section 1-3) indicate.

To explore the subjective wellbeing of beneficiaries we asked every respondent for their self-reported life satisfaction. For a sample of respondents, we asked more in-depth qualitative questions to understand the meaning of this score, and determine the factors that most commonly increase or decrease beneficiary wellbeing.

This data has the potential to inform moral weights, as it provides an alternate approach to informing the relative good of different types of outcomes. Specifically, we present analysis regarding the relative impact of doubling consumption to improving a respondent's health state, on their subjective wellbeing. As laid out by Plant (2018) this could form an alternate approach to GiveWell moral weights, in which the decision maker aims to maximize beneficiary wellbeing, rather than focusing directly on the trade-off between outcomes (averting death, and increasing consumption).

LIFE SATISFACTION

To capture the subjective well-being of respondents, we used Cantril's ladder (World Happiness Report 2018) to measure life satisfaction:

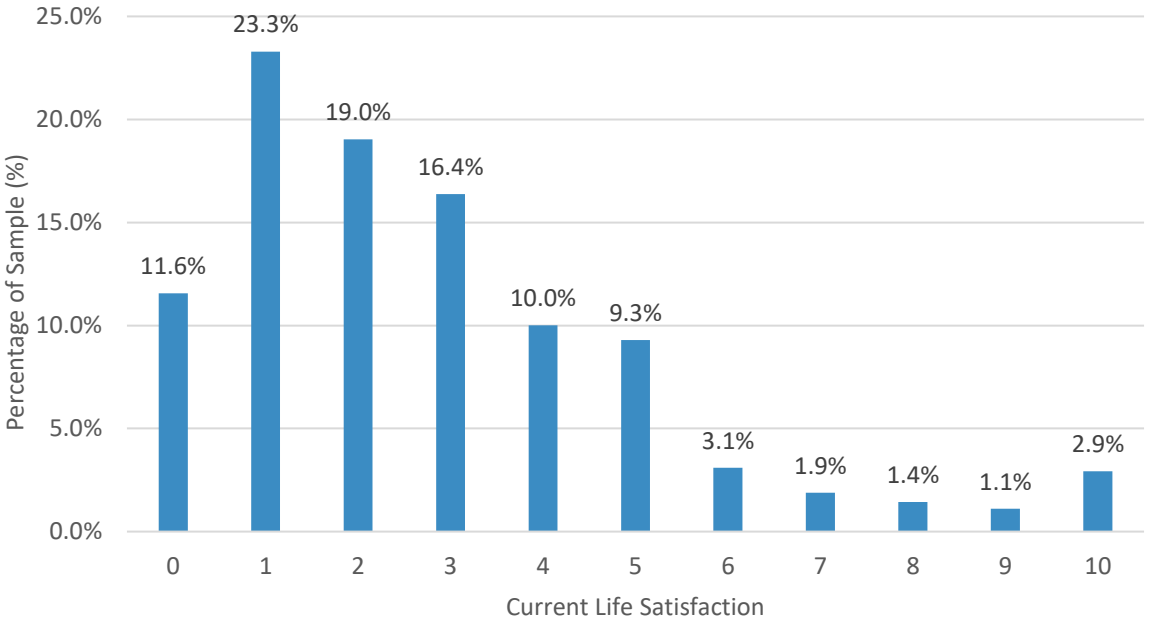
> *"Imagine a ladder from 0 to 10, and suppose that the top of it represents the best possible life for you. The bottom represents the worst possible life for you. On which step of the ladder would you say you stand at this point/today?"*

Using the self-reported life satisfaction as a reference point, we then asked about the factors influencing the life satisfaction scores, and the effect of different outcomes (increased consumption and averted deaths) on life satisfaction.

RESULTS

On the ladder from 0 to 10, respondents (n = 1808) report an average life satisfaction score of 2.8 (range 0 - 10, standard deviation = 2.32). This score is much lower than the 2019 World Happiness Report – Kenya had a score of 4.66 while Ghana had a score of 5.48.[68] This survey uses nationally representative samples, which means the average respondent would be richer than our purposefully sampled respondents from low-income regions and households. However, our results are still lower than expected based on extrapolation from other studies (see Appendix 6 for further exploration of this pattern).

Figure 7. Current Life Satisfaction Score Across the Full Sample.



Across Kenya and Ghana**, respondents describe their life satisfaction scores based on their current economic assets.** These include money, jobs, businesses, basic living needs (food, shelter, clothes and water) and education. Given how subjective life satisfaction scores are, the qualitative descriptions of each score vary from one respondent to the next. However, we noted the following patterns that align life satisfaction scores to economic assets:

- The most common reasons respondents gave for the low self-reported life satisfaction are lack of money (75% of respondents), lack of businesses/jobs (38% of respondents), lack of proper shelter (31% of respondents), lack of food (32% of respondents) and lack of education for respondents and their children (28% of respondents).[69]

[68] World Happiness Report 2019. https://ourworldindata.org/grapher/happiness-cantril-ladder
[69] In Appendix 6 we summarize all the reasons given for the self-reported life satisfaction scores, as well as their frequency across respondents.

- The things people mention they need to survive and to live a good life are also primarily economic. Of the 170 respondents surveyed in depth about their life satisfaction scores, 90% mention food, 61% mention shelter, 55% mention money, 43% mention clothes and water, 33% mention businesses/jobs and 30% mention good health as the most important for survival.
- One might expect factors such as health and overall wellbeing (joy, peace of mind etc.) to affect life satisfaction scores. However, most respondents associate a life satisfaction score of 10 with having a lot of money, cars, good houses, good education and good food. Conversely, they associate a life satisfaction score of 0 with a shortage or complete lack of these economic assets. Few (10%) mention factors such as health and overall wellbeing.

FACTORS ASSOCIATED WITH LIFE SATISFACTION

Across most of our respondents, we found a strong association between life satisfaction and their current living standards. Table 13, below, summarizes the results of regressing life satisfaction ladder on respondent characteristics in a sample including all 'poor' and 'wealthy' respondents.

Table 13. Regression of Life Satisfaction Ladder on Respondent Characteristics[70]

Variable	Coefficient (SE)
Log annual per capita con (USD, nominal)[71]	0.602*** (0.061)
Age	0.008** (0.004)
Female	0.261** (0.106)
Can read	0.211 (0.13)
Christian	0.014 (0.184)
Muslim	-0.384* (0.198)
Parent or grandparent to children in the household	-0.090 (0.132)
Self-reported long-term health condition	-0.331*** (0.108)
Urban	-0.006 (0.122)
Ghana	1.203*** (0.130)
N	2,032
R-squared	0.117

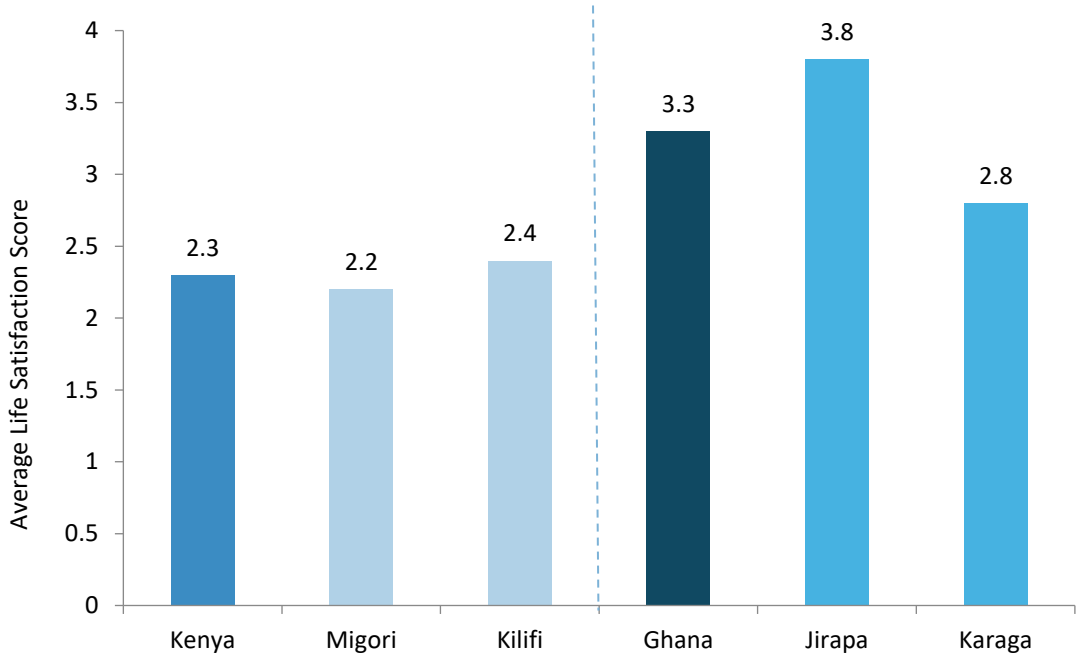
There are significant country-level and regional differences in current life satisfaction **—respondents in Ghana have a higher average life satisfaction score compared to respondents in Kenya.** As shown in Figure 8, respondents in Ghana, were on average 1 wrung higher on the life satisfaction ladder. This finding is consistent with the 2019 WHR results, where Ghana had a higher average life satisfaction score (5.48) than Kenya (4.66). There are significant region-level differences in life satisfaction

[70] This table presents coefficients from regressing responses to the life satisfaction ladder question on respondent characteristics. For the selection of regressors other than consumption or income, we follow Deaton and Kahneman (2010) as closely as possible. For region dummy variables, the omitted category is Migori (Kenya). The sample includes respondents classified as 'poor' as well as 'wealthy'. Standard errors are clustered at the sublocation level for Kenya and electoral area level for Ghana. The sample includes respondents classified as 'poor' as well as 'wealthy'. * denotes significance at 10%, ** 5%, and *** 1% levels.

[71] This coefficient implies that doubling consumption, which increases log consumption by 0.69, leads to an increase of 0.602*0.69=0.42 in the life satisfaction ladder. In contrast, Stevenson and Wolfers (2013) finds a lower coefficient of 0.25 among lower income countries.

scores—respondents in Jirapa (Ghana) have the highest average, followed by Karaga (Ghana), Kilifi (Kenya) and then Migori (Kenya).

Figure 8. Life Satisfaction Scores by Country and Region.



There is a positive and significant correlation between the wealth of respondents and their self-reported life satisfaction. **On average, respondents from wealthy households are on a higher step of the ladder than those from poor households.**[72] The results of the regression indicate that **doubling consumption corresponds to a 0.42 increase in the life satisfaction score.** This finding aligns with the reasons people gave for their reported life satisfaction—economic factors such as money and businesses/jobs were the main driving factors for the different ladder steps.

Like income, there is a positive and significant correlation (though lower) between the health status of respondents and their self-reported life satisfaction. **On average, respondents with self-reported long-term health conditions are on a lower step of the ladder than those without health conditions.** The difference between self-reported life satisfaction for people with health conditions and those without is statistically significant across respondents in Kenya and Ghana—on average, respondents with health conditions (n = 726) are on step 2.69 of the ladder while those without health conditions (n = 1322) are on step 3.12 of the ladder. The regression analysis effectively **implies that having a long-term condition moves you down the ladder by 0.33 steps.**

[72] We obtained the classifications of poor and wealthy households from two-staged sampling process: Participatory Wealth Ranking (PWR) and Progress out of Poverty (PPI) survey. In the PWR exercise, community leaders listed all the households in the village and collaboratively classify them as poor, moderate or wealthy. We then conducted a brief eligibility survey with all the poor households to identify those that were at or below our defined PPI threshold. Our 'wealthy' households had an average life satisfaction score of 4.3, while the 'poor' households had an average life satisfaction of 2.8.

Additional information on how different points on the ladder were described,[73] and how respondents expect different hypothetical scenarios to alter life satisfaction, is included in Appendix 6.

IMPLICATIONS

Overall, we find that life satisfaction is low across the low-income respondents in our sample. When asked what's missing in their life to achieve higher satisfaction, respondents almost all identified money, different material goods, or other economic factors. Plus, we find that consumption has the strongest association with life satisfaction out of all the factors considered. Having a self-reported long-term health consumption is associated with a smaller difference in life satisfaction compared with doubling consumption.

Taken together these findings suggest that the relative importance of increasing consumption relative to averting deaths might be higher than suggested by our preferences data. Staff members who think that maximizing subjective wellbeing is of high importance to moral weights (and that this is not adequately incorporated in beneficiary preferences) will likely put greater weight on the value of doubling consumption per person per capita, and less weight on averting deaths (across age groups). However, it is beyond the scope of this study to quantify this relationship. Larger datasets (ideally panel) from a relevant population would be required to give a more robust estimate of how subjective wellbeing varies with the two outcomes of interest to GiveWell (i.e. increasing consumption, and averting deaths).

[73] Including discussion with a number of respondents about whether a point on the ladder where life worse than death exists.

SECTION 5: ECONOMIC AND EMOTIONAL IMPACT OF DEATH

TAKEAWAYS

- If we consider a model of the net future economic contribution of an individual to their household, the relative value placed on individuals under 5 (compared to 5 and older) will be lower than if we rely on the results from Section 1-3 alone.
- Qualitatively, we identified substantial economic costs associated with the immediate aftermath of a death (primarily to cover funeral costs, and loss of income).
- We also found that the emotional impact of a death was generally higher for younger individuals. However, it is not possible to make a clear distinction in emotional impact between individuals under 5 and 5 and older.

This section details our findings on the estimated economic and emotional impact of death. It begins with an overview of our methods before covering the economic impact, then the emotional impact. We conclude with a discussion of the implications of our findings for GiveWell's moral weights.

OVERVIEW

An approach to assigning values to lives of people (particularly those in different age groups) is using their economic and emotional impact on the world around them. In particular, household members and communities benefit economically and emotionally from individuals, and such benefits would be lost as a result of an individual's death.

The economic contribution of different household members has previously been studied using panel survey data to examine how death results in changes in economic factors such as income and time-use.[74] However, few of these studies disaggregate the impact of death by age. Additionally, only a few studies report the economic impact of the death of a child, and all are conducted in HICs (Sanders 1980, Fox et al. 2014 and van den Berg et al. 2017).

As such, we conducted secondary analysis of recent household data from Kenya using the Kenya Integrated Household Budget Survey (KIHBS). We supplement this analysis with primary data from 47 qualitative interviews exploring the economic impact of death on household members and the community.[75,76]

[74] There are a number of studies on the economic impact of the death of a "prime-age" adult and parents in Africa. These include Ardington et al. 2012, Kenya MoH 2014, Deaton et al. 2008, Adhvaryu and Beegle 2014, Evans and Miguel 2007, De Weerdt et al. 2017, Beegle et al 2010, Yamano and Jayne 2004, Case and Menendez 2011, Beegle 2003 and Donovan et al. 2003. Most of these studies look at panel survey data and explain how the death results in changes in income and time use for family members.

[75] For the qualitative data, we first lifted the core ideas from respondent data; we then generated reproducible codes based on recurrent ideas and then categorized similar codes into broader themes

[76] We also explored collecting quantitative data on the economic contribution of different household members. However, there was no quick way of getting reliable data, and doing this on the full sample would be a time-consuming exercise. Given the existence of secondary data with a large sample size and that the main focus of our study is on beneficiary preferences, we focus on analyzing secondary data to answer these questions.

As there are few relevant and recent studies that quantify the emotional impact of death of a family member,[77] we piloted quantitative approaches to capture the length of grief and the number of people affected by death. We found that it was difficult for respondents to quantify their emotional experiences with precision, making it hard for us to capture this data robustly enough. We therefore decided to focus primarily on qualitative data to capture different people's emotional experiences with death across the four regions. In this section we present our findings from the 47 qualitative interviews on the emotional impact of death on household members.

ECONOMIC IMPACT OF DEATH

We first present findings from our analysis of secondary data that demonstrates the economic contribution of household members in order to infer, quantitatively, the long-term economic impact of their death. We then present qualitative data from our study on the economic impact of death on different household members.[78]

SECONDARY DATA FINDINGS

We analyzed publicly available secondary data from the Kenya Integrated Household Budget Survey (KIHBS) from 2015-2016.[79] We focus on low-income households, with average income per person per day of no more than 3 USD (PPP).[80,81]

We contrast our findings with the work of GiveWell staff member James Snowden, who developed models of household members' economic contribution and costs using existing studies in LMICs and intuition where data is lacking. Snowden's 2018 model is comprehensive and has been referenced by other GiveWell staff in their moral weights calculations.[82]

We broadly followed Snowden's approach, making the following three adjustments:

1. We incorporated the economic costs and contribution by age extracted from the KIHBS. This allows for a higher granularity of detail, as uniform assumptions about cost and contribution by age are not required. As a result of this change an individual's contribution to the household becomes net positive at a later age, and remains positive for longer.
2. To estimate net future contribution, we also incorporated present contribution while in the current age group. We felt the exclusion of this from the Snowden model was underestimating the economic contribution made by prime age adults whose current earnings are highest. This also results in an increase in the net future contribution of prime age adults.

[77] One of the few studies conducted on Sub-Saharan populations is Deaton et al. (2009). One study conducted in the US in 1980 used the Grief Experience Inventor (GEI) and MMPI to assess bereavement reactions (Sanders 1980). The study noted higher intensities of grief in parents surviving their child's death (relative to adults surviving a spouse or parent).
[78] Additional qualitative data on the economic contribution of different household members is in the attached "Beneficiary Profiles" supplementary document.
[79] For more details on this work and its methodology, see IDinsight memo "Informing Moral Weights", Section 3 "Economic Contribution Secondary Data Analysis: Findings for Kenya".
[80] This survey contains a sample of 21,773 households from 2400 clusters, drawn in a two-staged clustered sampling design. Our sample of low-income households include 16,801 households.
[81] 3 USD (PPP) is a widely used global threshold for poverty. We chose to use this threshold rather than the threshold for extreme poverty (1.9 USD PPP) so as not to systematically exclude poor households in urban areas, who are more likely to live above this absolute threshold.
[82] James Snowden's model for moral weights has been cited by GiveWell staff members Dan Brown, Isabel Arjmand, and Sophie Monahan.

3. We calculated the average value of an under-5 life and an over-5 life using the age distribution among poor households to be as comparable as possible to the results in Section 2. In contrast, Snowden used the distribution of malaria deaths by age to make the same calculation. As a result, our final ratio places lower relative value on individuals under 5 (as it is in this age group that a large proportion of malaria deaths occur).

Using the results of the updated model, and assuming a normal population distribution for the target population, we find a ratio of individuals under 5 to individuals 5 and older of 0.90. This is in contrast to the results of our methods capturing preferences (see Section 2), which find that individuals under 5 are valued higher at a ratio of between 1.2 from the individual perspective and 1.9 from the community perspective. However, adjusting the final assumption, (i.e. using the distribution of malaria deaths by age rather than the population distribution), results in a ratio of 1.1. Overall, we find that this approach results in lower relative value being placed on individuals under 5, but the exact ratio is highly sensitive to assumptions in the model.

Note that we compare our data and approach to Snowden's 2018 approach, which has already been updated since our work began. The step-by-step analysis approach, and figures directly comparing outputs are presented in Appendix 7.

PRIMARY DATA FINDINGS

In order to qualitatively understand the economic impacts of death of individuals on household members, we asked respondents about their experiences with death in the last 5 years.[83] Across the entire sample, we observed a high economic impact of the death of household members of all age groups. Similar to other studies, we observed that high funeral costs led to pronounced economic effects of death (Ardington et al. 2012 and Kenya Ministry of Health 2014). We present a heat map of all the themes that emerged in Appendix 7: Economic and Emotional Impact of Death.

The most common economic impacts of death, that were mentioned by more than a quarter of respondents, are:

- **Selling or renting household assets** such as land, livestock, trees and farm produce in order to get enough money for funeral costs.
- **Depleting all the household savings** or even working extra shifts in order to get enough money for funeral costs.
- **Receiving money, assets** (such as livestock and farm produce) **and services from community members** for funeral costs.
- **Spending time away from income-generating activities** during funeral activities and while grieving.

[83] If respondents indicated that they had lost household members within the last 5 years, we asked them to explain what 1) the economic effects of the death were and 2) the emotional effects were. We restricted the scope to 5 years so as to make it easier for respondents to recall the details. Due to the sensitivity around these kinds of questions, we asked general questions and let respondents go into as much detail as they wanted to. Enumerators were trained to probe only if the respondent expressed that they were comfortable talking about their experiences.

EMOTIONAL IMPACT OF DEATH

Individuals provide emotional value to those around them, and their death often has a tangible negative emotional impact on household members and others. Previously, the emotional impact of death has been quantified using panel data surveys that observe the same individuals and the effects of death on their lives over a period of time.[84] Since our study was not longitudinal, we captured, qualitatively, the self-reported impact on beneficiaries of the death of household members.[85] We describe the experiences that feature heavily across the sample below, along with key quotes from respondents.

FINDINGS

1. **For some, but not all, respondents' grief has had a tangible effect on their subjective wellbeing.**

Most respondents highlight the high emotional impact of death of various household members. For some, the impact manifests in their physical wellbeing. A respondent who lost his brother describes the emotional burden on their mother as follows:

> *"[Our] mother did not eat for a week after the death and was even almost dying herself" - Male, 30 Migori, Kenya.*

Another respondent describes the effect of the death of her father-in-law as follows:

> *"[The father of the deceased] fell sick due to the shock of the funeral, and he is still in the hospital (7 months after the death)" - Female, 49, Jirapa, Ghana.*

In contrast, some respondents note that the magnitude and burden of the grief of their loved ones was not too high. A respondent who lost his mother ascribes his low burden to his religious beliefs as follows:

> *"I was really sad because losing a parent is painful, but I was not too sad such that I couldn't work. The Bible also says that we shouldn't spend too much time worrying so I did not mourn for too long" – Male, 39, Kilifi, Kenya.*

As these quotes demonstrate, respondents experience varied levels of grief and for different periods of time. Importantly, they demonstrate that it's not easy to map out the magnitude of death across different household members because the experience varies from one person to another.

[84] Such studies include Evans and Miguel 2007, De Weerdt et al. 2017 and Sanders 1980.

[85] If respondents have experienced a household death in the last 5 years, we ask them to recall the emotional effects, including questions about whether it stopped members of the household from continuing their daily activities. Due to the sensitivity around these kinds of questions, we asked general questions and let respondents go into as much detail as they wanted to. Enumerators were trained to probe only if the respondent expressed that they were comfortable talking about their experiences. If the enumerator felt the respondent was uncomfortable discussing, or they refused to answer these questions, the interview was stopped (across all 47 interviews this happened on 2 occasions only).

2. The magnitude and experience of grief varies with factors such as age of the deceased, nature of the relationship to the deceased, and the health condition of the deceased.

Despite the highly subjective nature of grief, we observed variations in the magnitude and experience of grief depending on certain factors. Across most respondents, the most common factors were age of the deceased, the nature of the relationship to the deceased, and the health condition of the deceased. Similarly, previous studies reveal that the impact of death varies heavily depending the relationship of the deceased to the household members (Sanders 1980; Deaton et al. 2009) and whether the person was sick for a long period of time before dying (Donovan et al. 2003). We describe each of the factors we observed along with key quotes from respondents below.

a) In general, the emotional impact of death is higher for younger individuals

Across many respondents, the death of younger individuals has a high emotional impact. For most of these respondents, the grief experienced after the death of younger individuals is tied to the loss of current and future socio-economic potential. A respondent who lost two of his sons (one aged 8 and another aged 14) within a period of 4 years describes his experience as follows:

> *"I had hoped that [my 8-year old son] would grow up and become successful. All my joy was just taken away. I did not fully recover emotionally because barely 4 years later, another son [14-year old] died. Sometimes when I sit down somewhere, I still think about it and about if they were alive, they would be helping me around and I would be able to send them to the farm but now I have to do it myself" – Male, Migori, Kenya.*

Conversely, the death of older individuals causes a low emotional burden. A respondent who lost his father-in-law describes his emotional experience of the death as follows:

> *"There was no emotional impact because he was old. At that age, people are ok for him to die. It is worse when the person is young. He was using a walking stick so [he was] very old" – Male, 32, Karaga, Ghana.*

However, there are a few respondents who reported a high emotional burden upon the death of older individuals. A respondent who lost her 52-year old grandmother described her experience as follows:

> *"[I] did not work for three months after this death because [I] was really sad. [I] stayed at [my] parental home during these three months. During this time, [my] husband used to provide the money for the household's daily expenses" – Female, 19, Kilifi, Kenya.*

Prior to this study some GiveWell staff members theorized the magnitude of grief experienced by friends and family is highest for individuals between 5 years and 60 years, and lowest for individuals less than 1 year old.[86] Our overall findings align with GiveWell's framework, that the magnitude of grief is generally lower for older individuals (individuals above 60 years) than for younger individuals

[86] Based on assumptions in James Snowden's model, and the assumptions and descriptions detailed in Marinella Captiati's moral weights framework (see the notes within GvieWell's cost-effectiveness model for more details).

(individuals under 60 years). However, our findings also suggest that the emotional burden could be higher for individuals under 5 years than is currently assumed staff members due to their perceived future socio-economic potential.

b) In general, the emotional burden is higher for closer relatives

Most of the respondents with the highest reported magnitude of grief are those who experienced the death of close relatives such as children, parents or grandparents.[87] Conversely, those with lower magnitudes of grief lost distant relatives such as aunts, cousins, uncles, nephews and stepmothers. A family that lost one of their children describes their emotional burden as follows:

> *"…The grandparents of the child; and [me] (the father of the child) and mother were affected; [we] grieved for 3 months before [we] finally let go. [We] couldn't go to farm, and even lost some of [our] farm quotas and bambaram beans" – Male, 43, Jirapa, Ghana.*

Another respondent who lost his grandmother explained how sad he was after this death and how he did not work for two weeks because of how "[he] was fond of her" (Male, 21, Kilifi, Kenya). Even though the periods of grief are highly varied across respondents, our data reveals that people take longer time off work when they lose close relatives.

On the other hand, respondents who lost distant relatives did not report being too emotionally affected. A respondent noted that after she lost her cousin, she did not experience too much grief. She notes:

> *"Other than the cultural stipulations of not working before the deceased is buried, [I] was not impacted emotionally by this death such that [I] couldn't work" – Female, Migori, Kenya.*

Another respondent noted that after the death of his stepmother, he was not too sad that he could not resume work. "It was more about the [funeral] planning logistics" (Male, 35, Migori, Kenya). We observed such responses across respondents who had experienced the death of other distant relatives such as nephews, uncles, aunts and in-laws.

c) In general, the emotional burden is lower for sick individuals

Respondents who experienced the death of loved ones due to terminal illnesses reported having less periods and magnitudes of grief, perhaps due to the prior knowledge of the imminent death. A respondent who lost her cousin to cancer noted:

> *"I was sad but since I knew she had cancer, I was prepared for anything so I took 1 week before resuming normal activities" – Female, 51, Kilifi, Kenya.*

Another respondent experienced the death of a friend who was sick and who had been admitted at hospital at the time of death. The respondent (who was also old and sick) said that he did not grieve

[87] A quantitative assessment of bereavement in the US noted higher intensities of grief in parents surviving their child's death (relative to adults surviving a spouse or parent) (Sanders 1980).

too much because he could also imagine himself being in that position (Male, 65, Jirapa, Ghana). For these respondents, the emotional burden of death is low because of the awareness of the condition of the deceased. This finding is similar to a study in Rwanda that revealed that in contrast to deaths not due to illness, a majority of deaths due to illness indicated no effects on savings/debt, other income generating activities, diet and overall wellbeing (Donovan et a. 2003).

DISCUSSION

Overall, we find that:

1. **The net future economic contribution lost places lower relative value on individuals under 5 than** the preferences results contained in Section 2. However, we do not have high confidence in an exact value estimated through economic contribution approach as it is highly sensitive to assumptions used in the model.
2. **The impact on household consumption in the immediate aftermath of death is substantial**. A complete model of the economic and emotional impact of death should include this factor. However, we do not have clear data as to how these impacts vary with the age of the individual, nor do we have enough data to quantify an average value of this impact. We suggest that staff members who would like to incorporate this value into their moral weights conduct an additional review of the literature, or make assumptions based on the qualitative data presented in detail in Appendix 7.
3. **We also found evidence of high emotional impact of death, particularly for younger individuals.** We do not have enough data to conclude whether the emotional impact was lower for very young children than for individuals over 5.

There are a number of limitations to this aspect of the study. For the secondary data analysis, we only looked at data from the most recent Kenya household income and consumption survey so the quoted figures cannot be fully representative of the recipients of GiveWell's top charities. Similar surveys have been conducted in other GiveWell target countries, so this analysis could be replicated. While we do not expect this would drastically change the results, it could increase confidence in the robustness of the outputs.

Additionally, the primary data that we collected was purely qualitative, and was only captured in cases where the enumerator felt the respondent was comfortable discussing this sensitive topic. The results are therefore not comprehensive. Nonetheless, we hope that the data presented, along with the Beneficiary Profiles will give readers firsthand insight into the impact of death in these contexts that can inform their decision making.

CONCLUSION

Incorporating the preferences captured using the VSL and choice experiment methods in this study into GiveWell moral weights would result in:

1. Placing a higher value on averting deaths relative to doubling consumption,
2. Placing a higher value on averting the death of individuals under 5 than individuals 5 and older.

This would lead to a higher relative cost-effectiveness of charities whose good is achieved primarily by averting the death of young children (e.g. Helen Keller International, Malaria Consortium, and Against Malaria Foundation etc.).[88]

These findings are supported by qualitative data that suggests that methods were broadly well understood and that extreme valuations represented a clear moral stance for many respondents. This highlights that many respondents in low income settings may apply different ethical frameworks to the more utilitarian framework applied by many GiveWell staff members. These findings were further reinforced by secondary methods which also found a higher relative value for individuals under 5, and a similar range of values placed on life relative to consumption, despite applying different question framings (see Appendix 9).

However, there are a number of reasons why GiveWell may not want to completely default to the results of the preferences captured in this study. First, while our results as a whole are relatively reliable (within the limitations of a stated preference approach), and we believe the aggregated result gives a reasonable central estimate of true underlying preferences, there remains a high level of uncertainty in the methods we used. Obtaining a single estimate still requires a number of subjective judgements, including the best analytical approach, and the relative weight given to each of the methods. We believe it is important to recognize the value of the study — namely being one of the first studies to systematically estimate the preferences among aid beneficiaries over different outcomes — while accounting for its limitations when applying the results. We explore the methodological limitations further in Appendix 10.

Second, direct beneficiary preferences are just one possible approach to inform resource allocation decisions — GiveWell staff members may wish to weight other factors. In Section 4 and 5 we present data related to two potential alternate approaches. We find that placing more weight on maximizing wellbeing as opposed to satisfying preferences would likely result in a higher value being placed on income increasing interventions relative to life-saving interventions. We also find that relying on economic valuations of individual's contribution to their household will likely result in lower relative value being placed on individuals under 5. The decision about the relevance of these different approaches compared to preferences remains subjective.

[88] If the moral weights were changed to fully reflect the quantitative results of this study, the ranking of current GiveWell top charities would shift substantially and Helen Keller International, Malaria Consortium, and Against Malaria Foundation would be the most cost-effective charities (with charities such as Sightsavers and GiveDirectly becoming less cost-effective). It would also make it substantially easier for other charities whose main outcome is saving lives, particularly of children under 5, to reach and surpass GiveWell's cost-effectiveness threshold (2-3x as cost-effective as cash).

IMPLICATIONS BEYOND GIVEWELL

This study demonstrates that it is possible to inform the complex, moral trade-offs faced in development by capturing the preferences of the people affected by these decisions.

This study represents a substantial addition to existing VSL literature in which low income populations, particularly those in Sub-Saharan Africa, have previously been severely underrepresented. The data produced can immediately inform benefit-cost analysis (BCA) used in policy decision making in related populations. Incorporating this study's results into BCA offers two main advantages:

1. *Incorporating data from the countries of interest increases confidence in BCA output, as the underlying assumptions better reflect the preferences of target populations. This may increase the likelihood of decision makers using these results.*

   Current recommendations for BCA in low income countries rely on extrapolation of VSL figures from studies in the US, due to the lack of reliable data from LMICs.[89] This raises the concern that the values do not represent the true preferences of the population served. This study provides a reliable reference point from low income households in Kenya and Ghana, a population that is more similar to those targeted by BCA in most LMICs than those included in US VSL studies.

2. *Our data provides a more granular look at how VSL varies within countries, specifically among very low-income populations.*

   Current recommendations for BCA provide estimates for country-level VSL,[90] but many BCA studies focus on interventions that target a specific section of the population. This study provides insight as to how VSL varies within country for Kenya and Ghana that can be used to inform assumptions where BCA results are required at the subnational level.[91]

Beyond the VSL community, these results may also be used by foundations, and individual donors, to directly inform their giving and allocation decisions. More work is required to explore how to make these results accessible to more audiences, and to understand how organizations beyond GiveWell can best use this data.

[89] See Robinson et al (2019) guidance. While there are a small number of studies of VSL in LMICs, the majority have been based in China and South East Asia. These studies have informed recommendations about how to extrapolate from US figures to a lower income population based on GNI per capita. However, there is not enough primary data from LMICs of interest to come up with an estimate of VSL directly.

[90] Again, see Robinson et al (2019) guidance).

[91] For example, a recent study by IDSI conducted benefit-cost analysis to prioritise the use of different TB testing approaches. The underlying VSL values used were those estimated for the whole of South Africa based on recommendations, despite the fact that TB disproportionately impacts the bottom wealth quintile. The outputs of this analysis are highly sensitive to this assumption.

NEXT STEPS

This study fills a clear gap and can immediately inform GiveWell's moral weights and other BCA in low income countries. Our data focused on preferences related to one key trade-off, from a population of very low-income respondents that are of high relevance to GiveWell's decision. However, a much broader range of trade-offs, affecting many diverse populations, are faced by decision makers. As such, we believe this study should serve as the beginning of a larger push to capture and use beneficiary preferences across many areas to inform government and development sector resource allocation.

Some next steps to achieve this broader goal:

1. **Expand research agenda to capture:**
    a. **The preferences of populations across more diverse contexts**. Our sample was focused across 4 regions in Kenya and Ghana, and was not designed to be representative of these regions (as we were focussed on particularly poor households) or the population of these countries. There is a clear need for studies across more regions and countries to validate these results, and further explore variations in valuation between different populations.
    b. **Other types of preferences**. In line with GiveWell's priorities, this study focused on methods that directly capture the trade-off between increasing consumption and averting the deaths of individuals of different ages. However, there are many other beneficiary preferences that may be relevant to different decision makers. For example, our methods could be adapted to other non-health related outcomes that cannot be converted to dollars per lives saved. This could include preferences around the equality of distributions (i.e. should we target the poorest, or spread aid across larger populations), autonomy (i.e. should we prioritize public provision or interventions that increase beneficiary autonomy), and the relative value of health to other important outcomes (such as education or female empowerment).
2. **Work directly with different organisations to understand how others can incorporate preferences into their decision making.** This study was designed in partnership with GiveWell to inform the specific trade-offs they face, and the results presented to fit their decision-making system. By working with other organisations which face similar trade-offs we can better understand how others can use these results, and establish what additional data might be highest priority to collect.
3. **Develop tools that improve the accessibility of this data to a broader audience of policy makers and donors**. The resource allocation decisions addressed by this study, as well as the methods and analytical approach used, are complex. Simple tools that summarise the different ways to approach this problem, the best available literature, and the results of this study could make this decision-making process more accessible to a broader audience.

REFERENCES

Adhvaryu, Achyuta R., and Kathleen Beegle. *The long-run impacts of adult deaths on older household members in Tanzania*. The World Bank, 2009.

Alberini, Anna, et al. *Does the value of a statistical life vary with age and health status? Evidence from the United States and Canada*. No. 1318-2016-103522. 2002.

Alberini, Anna, et al. "Does the value of a statistical life vary with age and health status? Evidence from the US and Canada." *Journal of Environmental Economics and Management* 48.1 (2004): 769-792.

Ardington, Cally, et al. *The economic consequences of death in South Africa*. 2012.

Beegle, Kathleen. *Labor effects of adult mortality in Tanzanian households*. The World Bank, 2003.

Beegle, Kathleen, Joachim De Weerdt, and Stefan Dercon. "Orphanhood and human capital destruction: Is there persistence into adulthood?." *Demography* 47.1 (2010): 163-180.

Bénabou, Roland, Armin Falk, and Jean Tirole. *Eliciting Moral Preferences*. Mimeo, 2018.

Berti, Peter R. "Intrahousehold distribution of food: a review of the literature and discussion of the implications for food fortification programs." *Food and nutrition bulletin* 33.3_suppl2 (2012): S163-S169.

Case, Anne, and Alicia Menendez. "Requiescat in Pace? The consequences of high-priced funerals in South Africa." *Explorations in the Economics of Aging*. University of Chicago Press, (2011). 351-373.

De Weerdt, Joachim, Kathleen Beegle, and Stefan Dercon. "Orphanhood and self-esteem: an 18-year longitudinal study from an HIV-affected area in Tanzania." *JAIDS Journal of Acquired Immune Deficiency Syndromes* 76.3 (2017): 225-230.

Deaton, Angus. "Income, health, and well-being around the world: Evidence from the Gallup World Poll." *Journal of Economic perspectives* 22.2 (2008): 53-72.

Deaton, Angus, Jane Fortson, and Robert Tortora. *Life (evaluation), HIV/AIDS, and death in Africa*. No. w14637. National Bureau of Economic Research, 2009.

Donovan, Cynthia, et al. *Prime-age adult morbidity and mortality in rural Rwanda: Effects on household income, agricultural production, and food security strategies*. No. 1093-2016-87998. 2003.

Elias, Julio, Nicola Lacetera, and Mario Macis. "Efficiency-morality trade-offs in repugnant transactions: A choice experiment." (2016).

Evans, David K., and Edward Miguel. "Orphans and schooling in Africa: A longitudinal analysis." *Demography* 44.1 (2007): 35-57.

Fox, Melanie, Joanne Cacciatore, and Jeffrey R. Lacasse. "Child death in the United States: productivity and the economic burden of parental grief." *Death studies* 38.9 (2014): 597-602.

Garcia-Retamero, Rocio, Mirta Galesic, and Gerd Gigerenzer. "Do icon arrays help reduce denominator neglect?." *Medical Decision Making* 30.6 (2010): 672-684.

Garcia-Retamero, Rocio, Okan, Yasmina, and Cokely, Edward. “Using visual aids to improve communication of risks about health: A review.” *The Scientific World Journal*. (2012).

Hoffmann, Sandra, Qin, Ping, Krupnick, Alan, Badrakh, Burmaajav et al. “The willingness to pay for mortality risk reductions in Mongolia.” *Resource and Energy Economics* 34 (2012) 493-513.

Hoffmann, Sandra, Krupnick, Alan, and Qin, Ping. “Building a set of internationally comparable value of statistical life studies: Estimates of Chinese Willingness to Pay to Reduce Mortality Risk.” *Journal of Bnefit Cost Analysis* (2017) 8(2): 251-289.

Johansson-Stenman, Olof, and Peter Martinsson. "Are some lives more valuable? An ethical preferences approach." *Journal of Health Economics* 27.3 (2008): 739-752.

Johansson-Stenman, Olof, Minhaj Mahmud, and Peter Martinsson. "Does age matter for the value of life? Evidence from a choice experiment in rural Bangladesh." *Working Papers in Economics 389* (2009).

Kahneman, Daniel, and Angus Deaton. "High income improves evaluation of life but not emotional well-being." *Proceedings of the national academy of sciences* 107.38 (2010): 16489-16493.

Kenya Ministry of Health, Family Care International (FCI), International Center for Research on Women (ICRW), and the KEMRI/CDC Research and Public Health Collaboration. “A Price Too High to Bear: The Costs of Maternal Mortality to Families and Communities - Summary of Research Findings.” (2014)

Kremer, Michael, et al. "Spring cleaning: Rural water impacts, valuation, and property rights institutions." *The Quarterly Journal of Economics* 126.1 (2011): 145-205.

Lee, Minwoo, Sunhae Sul, and Hackjin Kim. "Social observation increases deontological judgments in moral dilemmas." *Evolution and Human Behavior* 39.6 (2018): 611-621.

León, Gianmarco, and Edward Miguel. "Risky transportation choices and the value of a statistical life." *American Economic Journal: Applied Economics* 9.1 (2017): 202-28.

Mahmud, Minhaj. “On the contingent valuation of mortality risk reduction in developing countries.” *Applied Economics* 41.2 (2011): 171-181.

Mobarak, Ahmed Mushfiq, et al. "Low demand for nontraditional cookstove technologies." *Proceedings of the National Academy of Sciences* 109.27 (2012): 10815-10820.

Nor, G., and M. M. F. Yusuf. "Value of life of Malaysian motorists: Estimates from a nationwide survey." *Journal of the Eastern Asia Society for Transportation Studies* 4 (2003): 275-284.

Plant, Michael. “A Happiness Manifesto: Why and How Effective Altruism Should Rethink its Approach to Maximising Human Welfare” *EA Forum* (2018)

Robinson, Lisa A., et al. "Understanding and improving the one- and three-times GDP per capita cost-effectiveness thresholds." *Health Policy and Planning* 32.1 (2016): 141-145.

Robinson, Lisa A., James K. Hammitt, and Lucy O’Keeffe. "Valuing mortality risk reductions in global benefit-cost analysis." *Journal of Benefit-Cost Analysis* 10.S1 (2019): 15-50.

Robinson Lisa A., et al. “Reference Case Guidelines for Benefit-Cost Analysis in Global Health and Development.” (2019)

Robinson, Lisa A., et al. "Valuing Children’s Fatality Risk Reductions." *Journal of Benefit-Cost Analysis* 10.2 (2019): 156-177.

Sachs, Jeffrey D., Richard Layard, and John F. Helliwell. *World Happiness Report 2018*. No. id: 12761. 2018.

Salomon, Joshua A., et al. "Common values in assessing health outcomes from disease and injury: disability weights measurement study for the Global Burden of Disease Study 2010." *The Lancet* 380.9859 (2012): 2129-2143.

Sanders, Catherine M. "A comparison of adult bereavement in the death of a spouse, child, and parent." *OMEGA-Journal of Death and Dying* 10.4 (1980): 303-322.

Shapiro, Jeremy. "Exploring Recipient Preferences and Allocation Mechanisms in the Distribution of Development Aid." *The World Bank Economic Review* (2019).

Shrestha, Maheshwor. *Get rich or die tryin': perceived earnings, perceived mortality rate and the value of a statistical life of potential work-migrants from Nepal*. The World Bank, 2017.

Stevenson, Betsey, and Justin Wolfers. "Subjective well-being and income: Is there any evidence of satiation?." *American Economic Review* 103.3 (2013): 598-604.

Tortora, Robert. "Sub-Saharan Africans rank the Millennium Development Goals (MDGs)." *Washington DC: Gallup World Poll* (2009).

van den Berg, Gerard J., Petter Lundborg, and Johan Vikström. "The economics of grief." *The Economic Journal* 127.604 (2017): 1794-1832.

Viscusi, W. Kip, and Clayton J. Masterman. "Income elasticities and global values of a statistical life." *Journal of Benefit-Cost Analysis* 8.2 (2017): 226-250.

Viscusi, W. Kip, and Joseph E. Aldy. "The value of a statistical life: a critical review of market estimates throughout the world." *Journal of risk and uncertainty* 27.1 (2003): 5-76.

Whittington, Dale, and Joseph Cook. "Valuing Changes in Time Use in Low-and Middle-Income Countries." *Journal of Benefit-Cost Analysis* 10.S1 (2019): 51-72.

Wilkinson, Thomas, et al. "Comparing the Application of CEA and BCA to Tuberculosis Control Interventions in South Africa." *Journal of Benefit-Cost Analysis* 10.S1 (2019): 132-153.

Yamano, Takashi, and Thomas S. Jayne. "Measuring the impacts of working-age adult mortality on small-scale farm households in Kenya." *World Development* 32.1 (2004): 91-119.

APPENDIX

APPENDIX 1: SAMPLE APPROACH AND DEMOGRAPHICS

SAMPLE APPROACH

Data collection was conducted in Kenya and Ghana between May and September of 2019. In order to identify a sample that represents a diverse set of viewpoints from potential beneficiaries, our sampling strategy was as follows:

1. Purposeful sampling at the county/district level, to ensure we conducted research in two geographically and culturally diverse areas with large numbers of typical beneficiaries.[92]
2. Random selection of sub-counties, sub-locations, villages, and eligible households to obtain a sample of typical beneficiaries within identified areas.[93]

In order to identify eligible households at the village level, we:

1. Conducted a Participatory Wealth Ranking (PWR) exercise:
   - Community leaders from all areas of the village were identified and invited to a central meeting,[94]
   - Together they defined the characteristics of poor, moderate, and wealthy households in their community,
   - We worked with leaders to list all households within their neighborhoods, and classified them as poor, moderate, and wealthy.
2. Conducted a brief eligibility survey including the Progress out of Poverty Index (PPI) with all households classified as 'poor'.
3. Randomly selected poor households from the shortlist of households that were at or below our defined PPI threshold.[95,96]

[92] Regions across Kenya and Ghana were ranked according to poverty levels and under 5 mortality. Any areas where data collection was deemed infeasible (for example, due to safety concerns) were excluded, as were areas in Kenya where Give Directly is currently scaling up operations (as we expected concurrent receipt or expectation of cash transfers would likely bias results). We chose the two highest ranked regions that offered geographic diversity (i.e. were as far apart as possible) and had a different predominant religion (in each country we aimed for one predominantly Christian, one predominantly Muslim region).

[93] The samples we obtained from each country are not representative of their regions and are not self-weighted since we did not use probability proportional to size sampling. To make them representative of poor households in the regions, we would need to weight the observations by the number of poor households in each region, which we lack.

[94] In the case where a village was too big for us to feasibly invite all leaders to this meeting (i.e. more than ~120 households) we first chose a random segment of the village.

[95] We chose thresholds for each country to balance the probability of type I and type II errors (i.e. we chose a threshold where the probability of including a household that is not poor is roughly equal to the probability of excluding a poor household).

[96] This combination of PWR and PPI-based eligibility verification has been used by Village Enterprise in Kenya and Uganda to identify "ultra-poor" households for their graduation program: https://villageenterprise.org/wp-content/uploads/2015/12/Participatory-Wealth-Ranking-and-PPIs_November_2015_AJ-white-paper.pdf. The main difference between our version and theirs is that in the interest of time we selected opinion leaders from the entire community by asking community leaders (e.g. village chairman or chief) rather than asking households from each part of the community for opinion leaders from their part of the community.

4. Randomly selected wealthy households from the list identified during PWR.[97]

We deviated from this protocol in a number of communities in Ghana. We found during piloting that communities were smaller, and more dispersed than in Kenya, making it more difficult to convene PWR meetings. Therefore, in communities we identified as small (<60 households) we completed PPI across the whole village and then randomly selected respondents from all eligible households, skipping the PWR stage.

Our target sample size was 450 poor respondents, and 75 relatively wealthy respondents in each region. As shown in Table 14, we achieved the desired sample in Kenya, but fell short of wealthy respondents in Ghana. This was due to challenges identifying wealthier households in both Karaga and Jirapa regions in Ghana; there was very limited variation in wealth within the communities, especially in rural areas. We frequently found that everyone in every village was eligible according to the PPI threshold. The wealthy households in our sample were therefore disproportionately found in urban areas.

Our consent rates for this survey were particularly high as we had visited the household the day before to conduct PPI, informing them about the study and letting respondents know we would return. There was some drop-off of respondents over the course of the interview; 94% completed all sections (see Table 14). In the case that an interview was not completed, we gained specific consent to use the data already captured. Therefore, data from the complete sections of incomplete interviews is still included in the analysis.

In all presented analysis, we report on poor respondents who have completed at least the VSL section (1820 poor respondents total), unless noted otherwise.[98]

Table 14. Summary of respondents approached, consented, and who completed the full interview across the four regions in Ghana and Kenya.

	Approached		Consented		Completed full interview	
	Wealthy	Poor	Wealthy	Poor	Wealthy	Poor
Migori, Kenya	78	476	78 (100%)	466 (98%)	74 (95%)	445 (93%)
Kilifi, Kenya	77	478	77 (100%)	468 (98%)	73 (95%)	452 (95%)
Karaga, Ghana	28	451	28 (100%)	451 (100%)	28 (100%)	440 (98%)
Jirapa Ghana	63	461	63 (100%)	461 (100%)	58 (92%)	419 (91%)
Total	246	1866	246 (100%)	1846 (99%)	233 (95%)	1756 (94%)

[97] We did not complete PPI with wealthy households because A) PPI, while sensitive to people close to the poverty line, does not distinguish well between moderate or well-off households as all are likely to receive a full score, B) we were confident in village leaders' ability to identify the wealthiest households in their community, and C) all households (poor and wealthy) completed the consumption module which gave us a good, comparable measure of income.

[98] Data from the households classified as "wealthy" is only used in subgroup analyses for VSL, choice experiments, and life satisfaction. This data is not included in the main estimates from our study, which focuses on poor respondents.

SAMPLE DEMOGRAPHICS

Table 15 presents descriptive statistics of all socio-demographic indicators across our respondents in Kenya and Ghana.

Table 15. Average values for all socio-demographic indicators for our full sample, and for our sample in Kenya and Ghana.[99]

Indicator	Aggregate Mean (SD)	Kenya Mean (SD)	Ghana Mean (SD)
Female[100]	54% (0.50)	54% (0.50)	54% (0.49)
Age	40.39 (14.98)	42.26 (16.02)	38.49 (13.60)
Literacy: can read[101]	48% (0.50)	75% (0.44)	21% (0.40)
Literacy: can write	44% (0.50)	68% (0.47)	19% (0.39)
Christian	61% (0.49)	79% (0.41)	42% (0.49)
Muslim	29% (0.46)	12% (0.32)	47% (0.50)
Household size	8.46 (5.89)	6.39 (3.11)	10.56 (7.16)
Number of children in the household	4.38 (3.57)	3.42 (2.49)	5.34 (4.19)
Number of children under 5	1.37 (1.63)	0.99 (1.09)	1.75 (1.96)
Number of children 5 or older	3.01 (2.61)	2.43 (2.07)	3.59 (2.94)
Households without children	8% (0.28)	11% (0.32)	5% (0.23)
Respondent is a parent to some children in the household	74% (0.44)	73% (0.44)	75% (0.43)
Has self-reported long-term health condition[102]	35% (0.48)	41% (0.49)	29% (0.45)
Has received any cash transfers	16% (0.37)	10% (0.30)	23% (0.42)
Has received any other charity assistance	35% (0.48)	50% (0.50)	20% (0.40)
Urban	21% (0.41)	20% (0.40)	22% (0.42)
Annual consumption per capita (nominal USD)	368.89 (2155.19)	310.91 (322.93)	428.49 (3051.44)
Annual consumption per capita (nominal USD, top 1% winsorized)[103]	296.73 (321.80)	307.93 (298.64)	285.22 (343.78)

[99] Total number of observations is 1820 for all indicators except "has self-reported long-term health condition", "has received cash transfers" and "has received any other charity assistance": n= 1807 and "annual consumption per capita" (both): n= 1805.

[100] We randomly selected whether a male or female adult respondent will be surveyed in a given household, and stuck to the gender assignment whenever possible. We set the selection probability to be 50-50 by default, but dynamically adjusted the probability to achieve an overall balanced sample.

[101] The literacy variables refer to being able to read or write Swahili for Kenya, and English or local language for Ghana.

[102] We asked only a yes/no question without asking for the condition, and only noted down the condition if they voluntarily mentioned it rather than probed, to avoid sensitivity around HIV. Hence our number may be an underestimate due to potential underreporting. Our number is unlikely to be an overestimate since all of those who said they had a long-term health condition gave information on the condition. The reported conditions vary in nature and severity.

[103] We found some outlier consumption values that are likely due to enumerator errors in data entry. To ensure the robustness of results we choose to use consumption per capita values censured at the top 1% in all analyses.

APPENDIX 2: INDIVIDUAL PERSPECTIVE: VALUE OF STATISTICAL LIFE (VSL)

In this appendix we provide an in-depth description of the approach we took to capture VSL, provide more data related to method credibility, and present the results of additional data analysis.

METHOD OVERVIEW

First, we trained respondents in basic understanding of probability and assessed their understanding with a set of test questions (all accompanied with visual aids):

1. **Basic understanding question 1:** *Imagine two lotteries. The chance of winning in one lottery is 5 in 1000, the chance of winning in the other lottery is 10 in 1000. Which lottery has the larger chance of winning?*
2. **Basic understanding question 2:** *Imagine two roads that are prone to accidents. The risk of dying in an accident on the first road is 1 in 1000, and on the second road is 3 in 1000. Which road is riskier?*
3. **Scale understanding question:** *Now imagine two different roads. The risk of dying in an accident on the first road is 1 in 100, and on the second road is 2 in 1000. Which road is riskier?*
4. **Basic understanding question 3:** *Imagine two people. The first person's chance of death is 5 in 1000 in the next 10 years. The second person's chance of death is 10 in 1000 in the next ten years. Which person is more likely to die in the next ten years?*
5. **Risk reduction question:** *Imagine a disease that kills 50 in 1000 people. There are three different vaccines available for the disease. Vaccine A reduces the risk of dying from this disease from 50 in 1000 to 20 in 10000, Vaccine B reduces the risk from 50 in 1000 to 40 in 1000, Vaccine C reduces the risk from 50 in 1000 to 30 in 1000.*[104]

Next, we captured respondent willingness-to-pay (WTP) for small risk reductions.

- **Scenario introduction, including baseline risk:** *Imagine a new disease that affects [ADULT/CHILDREN] in your village. The disease is rare so there is not much chance of you catching the disease. For every 1000 people, 20 will catch the disease in the next ten years. However, everyone who catches the disease will die. So, your risk of dying from the disease is 20 in 1000 over the next ten years (or 20 in 10,000 each year).*
- **Introduce first vaccine:** *A new [VACCINE/MEDICINE] has been made for the disease. It reduces your risk of dying from the disease form 20 in 1000 to [15 in 1000/10 in 1000] over the next ten years (or 20 in 10,000 each year to [15 in 10,000/10 in 10,000]). However, it is not available at the public health facility so you must buy it for yourself.*

[104] Most of these questions were adapted from similar studies of VSL in LMICs; Johansson-Stennman et al. (2009) used question 1,2, and 5 in their study in Bangladesh, and the papers by Hoffman et al. (2012, 2017) all include question 4. We adapted question 3 from studies of scale conceptualization in HICs (Garcia-Retamero et al. (2012)) in order to train respondents to consider the denominator in addition to the enumerator when presented with risk levels, and assess more complex understanding of probability.

- **Initial WTP:** *Do you want to buy this [VACCINE/MEDICINE] for [YOURSELF/YOUR CHILD]? If yes, how much would you be prepared to spend today to buy this [VACCINE/MEDICINE].* A payment card was used to help respondents to come-up with an acceptable figure.[105]
- **WTP in installments over 10 years:** *Now imagine that you are able to pay for this [VACCINE/MEDICINE] little by little over the next 10 years. How much are you willing-to-pay each month/year/total to receive this [VACCINE/MEDICINE], and so reduce your risk for the next 10 years?* A payment card was used to help respondents to come-up with an acceptable figure.
- **Repeat for second vaccine:** *A second new [VACCINE/MEDICINE] has been made for the disease. It reduces your risk of dying from the disease form 20 in 1000 to [15 in 1000/10 in 1000] over the next ten years (or 20 in 10,000 each year to [15 in 10,000/10 in 10,000]). However, it is not available at the public health facility so you must buy it for yourself.*
    - **WTP in installments over 10 years:** As above.
    - **Vaccine preference:** *If they were the same price, would you prefer to buy the first or second [VACCINE/MEDICINE].*

There were a number of randomized components in this section:

- **[ADULT/CHILD]:** Respondents who were main caretakers of children under 18 in the household were asked for their WTP for both themselves, and for a randomly selected child. Respondents who were not main caretakers of any children under 18 in the household were only asked for WTP for themselves.[106] The order in which the questions relating to adult/child appeared was randomized to account for any ordering effect.
- **[VACCINE/MEDICINE]:** The risk reducing item offered to the respondent was randomized such that half were asked WTP for a vaccine only, and the other half were asked WTP for a medicine only. This aimed to test for sensitivity to framing.[107]
- **[15 in 1000/10 in 1000]:** Respondents were offered two vaccines, offering either 5 in 1000 and 10 in 1000 risk reduction over ten years. The order in which these vaccines were presented was randomized such that half received 5 in 1000 first, and half received 10 in 1000 first. This allowed us to test for:
    - Population level scope sensitivity in WTP, by testing if respondents were on average WTP more for a higher risk reduction (external scope test),

[105] A payment card is a piece of paper (or a computer screen) with various monetary values, and respondents will choose the highest value they are willing to pay from the options (i.e. the highest value on the card not exceeding their willingness to pay). This instrument has been used in many VSL studies, and is the favored means to capture WTP in the Hoffman et al. studies that have been replicated in multiple countries including China (2017) and Mongolia (2013). In our case, we do not restrict choices to be from the payment card, but merely use the payment card as a tool to help respondents come up with values.

[106] For households with both children under 5 and those 5 and older, we randomly selected one of the two age groups, and then randomly selected a child from that age group to be in the child VSL questions. We dynamically adjusted the probability to achieve an overall balanced sample.

[107] We stopped including the medicine framing in Ghana as during pilot we found 1) there was no clear distinction between the two in the local languages where we are working, 2) when the distinction is clarified many respondents thought medicine would always completely cure the disease.

- Individual level scope sensitivity in WTP, by testing if respondents offered more for the more effective vaccine (internal scope test).

METHOD VALIDITY

We found very low levels of scenario rejection (less than 2% of respondents refused to offer any money for the hypothetical vaccine/medicine).

Understanding small probabilities

Respondents demonstrated reasonable understanding of small probability. In Table 16 we include an overview of respondent small probability understanding and in Table 17 we present the results of our specific understanding questions.

Table 16. Results of the small probability and risk reduction questions.[108]

Result	Aggregate	Kenya	Ghana
Answered all 4 probability questions correct the first time	58%	45%	71%
Required at most 1 additional explanation to answer all 4 probability questions correctly	22%	29%	16%
Required 2 or more additional explanations for any of the 4 probability questions	20%	26%	13%

Table 17. Summary of respondent understanding of small probability questions.

Understanding question	% correct first time
Basic understanding question 1 (two lotteries)	81%
Basic understanding question 2 (two roads)	93%
Basic understanding question 3 (two people)	84%
Risk reduction question 4 (three vaccines)	84 %
Scale understanding question (two roads, different denominators)	34%

For the basic understanding questions, our respondents performed better than those in a study in Bangladesh with similar questions (Mahmud, 2011).[109] Additionally, our respondents performed

[108] See "Method Overview" section above for the full list of the small probability and risk reduction questions. Note that we do not include the scale question in our analysis of basic understanding because our expectation for respondents answering this question correctly was low.

[109] 74% answered the "two lotteries" question correctly (81% of our respondents did), and 83% for the "two roads" one (93% of ours did). See Table 2 of https://www.tandfonline.com/doi/abs/10.1080/00036840600994252.

similarly as those from a study in China (Hoffman et al., 2017)[110] and a study in Malaysia (Ghani and Yusoff, 2003).[111] They performed worse than respondents of a study in Mongolia (Hoffman et al., 2012).[112] Compared to stated preference VSL studies from high-income countries, our respondents perform slightly worse than respondents when asked the "two people" question (basic understanding question 3 above). In a study in the US and Canada (Alberini et al., 2002), 88% of respondents got this question right the first time,[113] whereas 83.7% of our sample did.

For the scale understanding question comparing 1/100 and 2/1000, 33.8% got it right the first time, and (including this 33.8%) overall 91.5% got it right with at most 2 explanations. There are a number of studies on the phenomenon of denominator neglect.[114] Garcia-Retamero, Galesic, and Gigerenzer (2010) conducted experiments in Germany and found that in some cases, up to 50% of respondents compared risk reductions incorrectly due to denominator neglect, and that using visual aids (which we used in our scale question) significantly reduced the proportion who misunderstood.[115] Note that this question is not part of our criteria for sufficient basic understanding of probabilities.

Validity Tests

To assess the validity of our results we conducted three main tests, summarized in Table 18. First, we look at the representativeness of the sample; we drop a number of respondents from our estimation model due to poor understanding, so we test for bias in the type of respondents that are dropped. Next, we look at the scope sensitivity of our respondents. This examines whether WTP is proportional to the level of risk reduction presented. We test internal scope, which looks at whether each individual respondent pays more for a higher risk reduction; paying less for a higher risk reduction demonstrates low understanding of the scenario. Next, we test external scope, which looks across the full sample to test if respondents that are randomly presented the high-risk reduction first, pay more than those that are randomly presented the low risk reduction first. Finally, we look at the construct validity, which tests how VSL is correlated with key respondent covariates.

Overall, we find good evidence of method validity. A small proportion (11.6%) of our respondents failed the internal scope test, and were dropped from our final estimation. Once the least stringent restrictions on understanding were applied, our sample passed the weak external scope test which is seen externally as the most important threshold for internal validity of results. We also passed the construct validity test; WTP was significantly correlated with respondent income.

[110] Overall, 81% passed the probability understanding tests ("FLAG1"). It is unclear what the tests are. See Table 3 of https://www.cambridge.org/core/journals/journal-of-benefit-cost-analysis/article/building-a-set-of-internationally-comparable-value-of-statistical-life-studies-estimates-of-chinese-willingness-to-pay-to-reduce-mortality-risk/F0736C663CC8C30263752C6E7D187526.

[111] 93% answered the "two roads" question correctly (93% of ours did). See the "Results" section of https://www.semanticscholar.org/paper/VALUE-OF-LIFE-OF-MALAYSIAN-MOTORISTS%3A-ESTIMATES-A-Ghani-Yusoff/4d2e660287a2889947029536470253 8b6df0747d.

[112] 99% of respondents answered the "two people" question correctly (84% of ours did). See Table 2 of https://www.sciencedirect.com/science/article/abs/pii/S0928765512000218.

[113] See Table 4, and the "C. Probability Comprehension and Acceptance of the Scenario" subsection of the "IV. Sample Characteristics and Responses", https://www.sciencedirect.com/science/article/abs/pii/S0095069603001232.

[114] For a review, see https://www.questia.com/read/1P3-3931554511/the-denominator-neglect-in-decision-making.

[115] See Figure 2 of https://pure.mpg.de/rest/items/item_2099684/component/file_2099683/content.

Table 18. Summary of VSL validity tests.[116]

Type of Validity	Test	Result
Representativeness of sample	Balance in demographics of VSL estimation sample and full sample	We use subsample 'Model A' of the adult VSL sample (63% of the latter), as per our pre-analysis plan.[117] There are significant differences in some demographic variables between it and the full sample (see Table 19). As a robustness check, we also estimated VSL using inverse probability weighting (IPW) to adjust for the probability of being included in the estimation subsample as a function of demographics.
Scope Sensitivity	Internal scope test	1. **11.6% of the entire sample strictly fail**, that is, the same respondent pays less (more) for the higher (lower) risk reduction. 2. **70.0% of the entire sample strictly pass**, that is, pay less (more) for the lower (higher) risk reduction. 3. 18.4% of the entire sample pay the same for both risk reductions. [118] 4. In a t-test comparing WTP for the two risk reduction levels (answered by the same respondents) in the overall sample, the average value for the higher risk reduction is significantly higher (p-value <0.001).[119]
	Weak external scope test	**Passed with `Model A':** among `Model A', on average, the group of respondents who are given the higher risk reduction for the first WTP question report higher willingness to pay.[120]
	Strong external scope test	**Failed:** On average, the group of respondents given higher risk reductions for the first WTP question report higher willingness to pay, but the magnitudes are not proportional to the risk reduction levels.[121]
Construct Validity	WTP variation across key covariates	In both Kenya and Ghana, log per capita consumption is positively correlated with WTP values.

[116] This table describes results focusing on adult VSL. We include footnotes on results for child VSL whenever applicable.
[117] We pre-specified three potential subsamples for estimation of VSL values based on different requirements of respondent understanding, and that we would take the main estimate from the least stringent model that passes the weak external scope test if any of them passes, and if not, from the most stringent model. Since we passed the weak external scope test, we used model A in which respondents who fail the internal scope test, or need two or more explanations for any of the probability test questions, are dropped (model B and C have more stringent requirements). For more details, see our pre-analysis plan.
[118] For child VSL, these numbers are: 10.7%, 70.6%, and 18.7%.
[119] This also applies to child VSL.
[120] For adult VSL, for the first risk reduction level given to respondents, the average WTP among Model A respondents is 178.7 USD for those first given 5/1000 (SE: 14.5), 261.9 for those first given 10/1000 (SE: 19.0), and the difference is statistically significant at the 5% level (p-value 0.001), i.e. passing our pre-specified threshold for the weak external scope test. For child VSL's 'Model A' estimation subsample, this difference is also statistically significant at the 5% level (p-value 0.006).
[121] This is failed by all estimation models (subsamples with different requirements of understanding) for adult and child VSL. Passing this test would require the average WTP for the first 5/1000 risk reduction and half of the average for the first 10/1000 risk reduction be statistically indistinguishable.

Table 19. Comparing respondent characteristics between the analytical sample and the full sample.

Variable	Adult VSL Model A Sample (SD)	Rest of Adult VSL Sample (SD)	Difference (SE)
Female	0.522 (0.500)	0.561 (0.497)	-0.039 (0.024)
Age	39.583 (14.483)	41.478 (15.663)	-1.896*** (0.731)
Can read	0.475 (0.500)	0.474 (0.500)	0.002 (0.024)
Christian	0.566 (0.496)	0.670 (0.471)	-0.104*** (0.024)
Muslim	0.336 (0.472)	0.229 (0.421)	0.106*** (0.022)
Household size	8.938 (6.143)	7.784 (5.429)	1.154*** (0.288)
Number of children	4.633 (3.641)	3.985 (3.402)	0.648*** (0.174)
Is parent to some children in household	0.758 (0.429)	0.722 (0.448)	0.035 (0.021)
Has self-reported long-term health condition	0.352 (0.478)	0.351 (0.478)	0.001 (0.023)
Urban	0.242 (0.429)	0.163 (0.370)	0.080*** (0.020)
Annual consumption per capita (nominal USD, top 1% winsorized)	300.059 (332.095)	290.946 (307.251)	9.113 (15.879)
N	1,126	663	1,820

This table presents demographics of all adults in the adult VSL sample and those in the adult VSL estimation subsample ('Model A'), as well as differences between the two groups. * denotes significance at 10%, ** 5%, and *** 1% levels, for t-tests between values in the two groups

Qualitative Assessment of Understanding

We also examined our qualitative data, to see what evidence we have to support that respondents largely understood the presented scenarios, and to identify any misunderstandings or misconceptions that may have affected our data. Overall, a majority of respondents interpreted our questions correctly. As shown in Table 20 we found evidence of respondents clearly comparing and considering the risk reduction levels, and mapping WTP to the value of different lives. However, there were some common pitfalls that may have biased our VSL results. In particular, we know that for some we failed to overcome the liquidity constraint and the respondents answered purely based on available income. A number of respondents also anchored their response to the market value of vaccines/medicines. This is of some concern for the relative value of children vs adults, as a number of respondents noted that they would pay less as children's medication is typically cheaper.

Table 20. Qualitative data from focus groups exploring respondent understanding of VSL.

I. Evidence of good understanding	
Correct insight inferred	**Direct quotes from focus group discussions**
Respondents were comparing multiple levels of risk reduction when giving WTP values	*"I paid more for the one that saves 10 people because it saves more people and the other one saves less people. For instance, when you are driving a car, you'd rather run into one person that into many people"*
Respondents were mapping WTP to value of lives	*"I paid the same amount [for children and adults] because all of us are human beings and there is no reason, I should pay more for one. We both have souls and hearts so it doesn't matter if one is a child or an adult"*
Respondents understood that the WTP installment options were removing their current credit constraints	*"Paid same amount because of lack of money, but he would be willing to pay a lot more in installments"*
II. Evidence of misunderstandings	
Type of misunderstanding	**Direct quotes from focus group discussions**
Respondents failed to treat each vaccine scenario as new and unrelated to previous ones	*The respondent paid less because he didn't think he could afford to spend a lot for another vaccine; but he didn't really take into consideration the lesser people that would be saved*
Respondents believed that stating WTP was setting the price and therefore moderated their response to keep the vaccine affordable	*"…I saw that it is better if I put the same amount because I thought that if I pay more for one and not the other, then it will be expensive even for the one buying the medicine, so I was thinking a lot about the cost to the one buying the medicine."*
Respondents WTP was still limited by liquidity, even with small installments	*"I gave amounts based on what my current income is, so I gave the same for both"*
Respondents failed to understand risk reduction probabilities	*She didn't know what quantity she needed to buy to heal her fast, so she just paid the same amount [for both risk reductions]*

Mean vs. Median VSL

We take the mean VSL values to be consistent with the LMIC stated preference VSL literature. In stated preference VSL studies done in the US, medians are often reported, which we also report here for completeness. The median VSL for 5/1000 risk reduction among "Model A" respondents is about 19,420 USD, much lower than the mean of 41,680 USD. This pattern is consistent with some VSL studies in high income countries, e.g. Alberini et al. (2004) in the US and Canada,[122] which reports a mean WTP value of 770 USD (standard deviation: 2049 USD) and median WTP value of 350 USD for a 5/1000 risk reduction for the US sample, and a mean WTP value of 466 USD (standard deviation: 834 USD) and median WTP value of 263 USD for a 5/1000 risk reduction for the Canada sample. Although in this study and many other VSL studies in the US the mean and median values are calculated using

[122] Table 6 of https://www.sciencedirect.com/science/article/abs/pii/S0095069603001232

distributional assumptions.[123] The EPA uses both mean and median VSL values from stated preference studies, where the mean values are often calculated.[124] Even though the median value is much lower than the mean in our case, we believe it is reasonable to use the mean value from a sample that exhibits reasonable understanding of the questions in order to capture the average value in the population.

Liquidity constraint

While we do not have a way to completely eliminate the concern that liquidity constraint affects responses, we did find that WTP increases substantially with the small installment repayment reframing and that the resulting values are not high relative to respondents' income. The average initial WTP for the 5/1000 vaccine is 11.8 USD (SD: 69.3 USD) rising to 208.4 USD (SD: 459.7 USD) when offered to pay in small installments. This suggests offering small installments results in an average WTP that's 118.2 times higher than asking for a hypothetical immediate payment. Further, for 95% of respondents, the total WTP under the small installment repayment framing is greater than the initial WTP. This does indicate that the framing has an effect in relaxing the liquidity constraint even if it is not entirely overcome. On average, with the small installment framing respondents are WTP 11.7% of monthly consumption per capita, and 1.7% of monthly total household consumption. Responses on how respondents would come up with the money are about evenly distributed among multiple options (savings, selling assets, working more etc.), implying that not many respondents had enough cash on hand to repay every period.

Vaccine-medicine framing effect

We used the Kenya sample to test whether small changes in framing result in substantially different values.[125] 477 Kenya respondents received the medicine framing and 438 received the vaccine one. For adults, medicine has an average WTP of 233.9 USD, and vaccine 250.1 USD, but the difference is not statistically significant (p-value of t-test: 0.61); for child, average WTP is 194.2 USD for medicine and 238 for vaccine, with the difference being not statistically significant (p-value of t-test: 0.30). Overall, we did not find any evidence of framing effect in this study.

Order effect: We randomize the order of asking about the respondent themselves and the child. Asking about the adult first increases both adult and child VSL values, though only the former effect is statistically significant (p-value of t-test: 0.01). We average across the two orderings for the final estimate to remove the influence of the order.

[123] Many of these studies use dichotomous choice, where respondents are given a number of values and state whether they are willing to pay the value to reduce mortality risk by some amount. They often assume the true values follow a Weibull distribution whose parameters are estimated using maximum likelihood, and calculate the mean and median values accordingly.

[124] https://yosemite.epa.gov/sab/sabproduct.nsf/0/0CA9E925C9A702F285257F380050C842/$File/VSL%20white%20paper_final_020516.pdf

[125] In Ghana, we found that there was limited difference between a vaccine and medicine when translated to the local languages (i.e. they had the same word, and more explanation was required to make the difference clear). Therefore, we did not randomise the framing – we just used 'vaccine' throughout.

ADDITIONAL ANALYSIS

Table 21 below presents results from a construct validity test: a regression of the willingness to pay (WTP in USD) to reduce own mortality (from the first risk reduction level asked, with repayment in small installments) on respondent characteristics across the full sample. Across all respondents, there is significant correlation between willingness to pay and gender, age, region and consumption. Being female is associated with a $82.47 decrease in WTP and being a year older is associated with a $1.67 decrease in WTP.

Table 21. Regression of willingness to pay on respondent characteristics for adult VSL.

Variables	Coefficient (SE)
Risk reduction: 5/1000	-25.62 (17.07)
Female	-82.47*** (23.26)
Age	-1.67** (0.67)
Urban	35.51 (23.97)
Can read	1.73 (22.83)
Christian	-20.78 (51.41)
Muslim	-104.7* (59.31)
Household size	2.11 (1.89)
Self-reported long-term health condition	-21.05 (22.82)
Satisfaction ladder	6.41* (3.46)
Karaga (Ghana)	141.20** (55.87)
Jirapa (Ghana)	-56.49** (22.99)
Kilifi (Kenya)	122.50*** (33.97)
Log annual consumption per capita (nominal USD)	49.22*** (11.47)
N	2,004
R-squared	0.049

Log consumption per capita is constructed from consumption values with the top 1% winsorized. For region dummy variables, the omitted category is Migori (Kenya). The sample includes respondents classified as 'poor' as well as 'wealthy'. Standard errors are clustered at the sub-location level for Kenya and electoral area level for Ghana. * denotes significance at 10%, ** 5%, and *** 1% levels.

Below we illustrate how VSL is distributed across the different age groups. Note that the sample is small when we look into each age group (particularly for children), and so can be easily skewed by individual high values (particularly for children; see values for age 7 & 8 below). As a result, it is not possible to draw any clear conclusions from these distributions.

Figure 9. Child VSL (reported by caretakers) by child age (0-18 years).

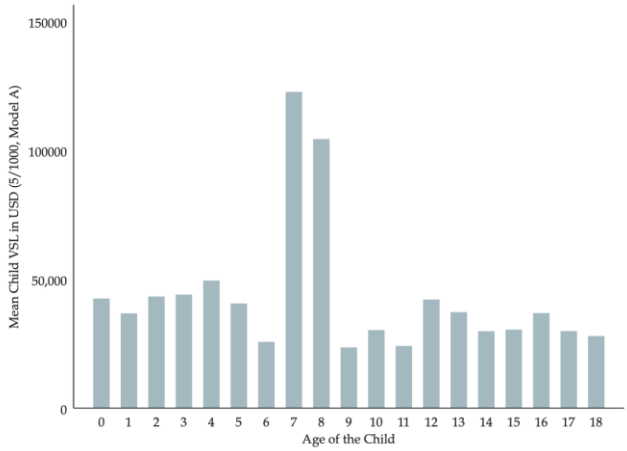
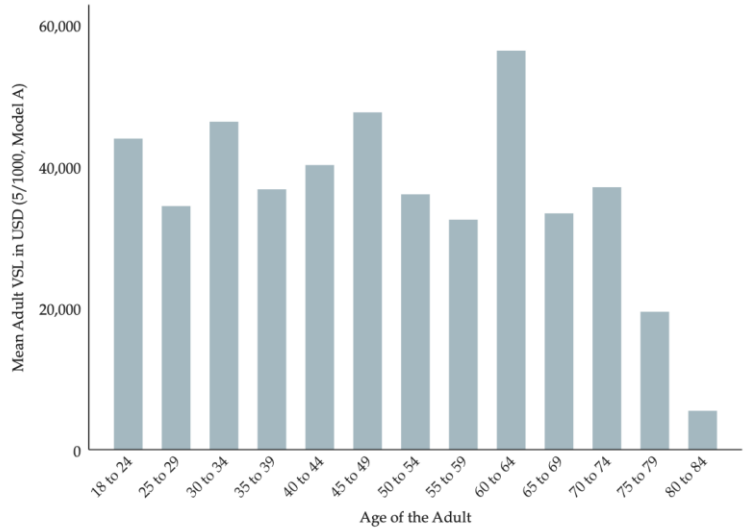


Figure 10. Adult VSL by age (18-84 years).



Finally, below we demonstrate the ranges across the adult VSL estimates across the different potential estimation models.

Table 22. Ranges across adult VSL estimates.

Model	Aggregate Mean in USD (SE); N	Kenya Mean in USD (SE); N	Ghana Mean in USD (SE); N
VSL: 5/1000 (Full sample)	41671 (3076); N= 893	47354 (4397); N = 470	35356 (4263); N = 423
VSL: 5/1000 (Model A)	35733 (2909); N= 544	41125 (3900); N = 247	31248 (4216); N = 297
VSL 10/1000 (Full sample)	24414 (1474); N= 896	24868 (2270); N = 420	24012 (1922); N = 476
VSL: 10/1000 (Model A)	26194 (1895); N = 582	26398 (3157); N = 231	26060 (2361); N = 351

As shown in Table 22, the central estimate varies depending on the model used.[126] In line with previous literature, and our pre-analysis plan, we chose to use the VSL from the risk reduction that's most relevant to GiveWell (5/1000) for our central estimate. Furthermore, we estimated VSL using inverse probability weighting to adjust for difference in demographics between our estimation subsample ('Model A') and the full sample (as discussed in Table 8 above).[127] For the aggregate sample across Ghana and Kenya, the mean weighted adult VSL is 35,999.69 USD (SE: 2970.26).

[126] Table 22 applies to VSL ranges for adults only. We, however, found similar ranges around the central estimate across other age groups.

[127] Specifically, we estimate a probit model of the probability of being included in `Model A' among adult VSL respondents as a function of respondent characteristics (gender, age, literacy, religion, household size, number of children, whether they are a parent to a child in the household, whether they reported having a long-term health condition, and winsorized annual consumption per capita), and used the inverse of the predicted probabilities as weights in the estimation. The generalized method of moments (GMM) was used to obtain consistent standard errors.

APPENDIX 3: COMMUNITY PERSPECTIVE: MONETARY VALUE OF LIFE

In this appendix we provide an in-depth description of the approach we took to capture the monetary value of life from the community perspective (with a choice experiment), provide more data related to method credibility, and present the results of additional data analysis (subgroup analysis and switching patterns).

METHOD OVERVIEW

First, we captured respondent perception of cash transfers, and prompted respondents to think about the impact of cash transfers, by asking how they, and then how most poor households, would use a $1,000 cash transfer.

Next, respondents were asked to imagine they were the decision maker for their community and that they must choose between two potential programs (that both cost the same) available to their community.

Respondents were first asked to make a choice where one program is clearly dominant, in order to assess basic understanding of the scenario:

"Program A saves the lives of 6 children aged 0-5 years AND gives $1,000 cash transfers to 10 families. Program B saves the lives of 5 children aged 0-5 years AND gives $1,000 cash transfers to 6 families. Which one would you choose?"

Respondents were then asked to make 3 similar choices; all attributes were kept consistent except for the number of cash transfers provided by Program B (which is randomly selected from a distribution as highlighted in blue). Respondents were asked to explain their reasoning after the first choice was presented.

"Program A saves the lives of 6 children aged 0-5 years AND gives $1,000 cash transfers to 5 families. Program B saves the lives of 5 children aged 0-5 years AND gives $1,000 cash transfers to [15/25/35/45/55/65/75/85/95/105] families. Which one would you choose?"

Enumerators used a visual aid to demonstrate what each program offers, which they used to highlight the difference between the two programs (for example, by pointing out that Program A saves 1 more life, while Program B provides 20 more cash transfers).

If the respondent consistently preferred one program, we presented additional 'extreme choices' to test for respondent switching:

- If they consistently preferred the program with more cash transfers, we reduced the number of cash transfers to 6 (effectively comparing 1 life to 1 cash transfer).
- If they consistently preferred the program that saves more lives, we increased the number of cash transfers to 1000, and if they still didn't switch, to 10,000.
- If they still didn't switch at these extreme levels, the respondent was asked to explain why.

We also looked at consistency of preferences within the initial three choices. For example, if a respondent preferred the program with more cash transfers, but then switched to the program that

saved more lives when the number of cash transfers was increased, we classified the individual as inconsistent.

We only keep respondents who both passed the dominance test and exhibited no inconsistencies in their choices in the estimation sample.

PERCEPTIONS OF CASH TRANSFERS

As outlined in the method overview section above, we asked respondents what they themselves and poor households in the community would do with a hypothetical 1,000 USD cash transfer. The question was open-ended and we recorded responses given by each respondent.

For what they would do themselves, 30.6% mentioned consumption-related categories (food, clothes, alcohol, entertainment, or religious spending), 93.5% mentioned investment-related categories (children's education, business, home improvement, livestock, land, or farm inputs), and 0.2% mentioned negative categories (defined as relating to alcohol, squandering, wasting, or gambling).

For what other poor households would do, 44.3% mentioned consumption-related categories, 86.3% mentioned investment-related categories, and 6.5% mentioned negative categories (alcohol, squandering, wasting, or gambling).

Overall, we found minimal evidence of negative perceptions of cash transfers that might bias responses to our choice experiment.

METHOD VALIDITY

A large proportion of respondents (95.3%) passed our basic scenario understanding test, and 92.1% demonstrated consistent preferences across their choice set. As a result, few observations were dropped for our full estimation sample (11.6%).

Table 23. Summary of Validity tests for cash vs life choice experiment.

Test	Result
Dominance Test	95.3% of respondents picked the correct choice when presented with a dominance test choice.
Consistency Test	92.1% of respondents were consistent across the different choices presented to them.
Both tests	88.4% of respondents passed both tests.
Sample balance test	There were few statistically significant demographic differences between the sub-samples that passed the dominance and consistency tests, and the full sample (see Table 24 below).

Table 24 below compares respondent characteristics between respondents in our estimation sample (those who passed the dominance and consistency test) and the rest of the sample.

Table 24. Comparing respondent characteristics between the estimation sample and the full sample.[128]

Variable	Cash vs. Life Estimation Sample (SD)	Rest of Sample (SD)	Difference (SE)
Female	0.52 (0.50)	0.61 (0.49)	-0.09** (0.04)
Age	40.14 (14.86)	42.11 (15.94)	-1.97* (1.10)
Can read	0.48 (0.50)	0.43 (0.50)	0.05 (0.04)
Christian	0.60 (0.49)	0.63 (0.48)	-0.03 (0.04)
Muslim	0.30 (0.46)	0.22 (0.42)	0.08** (0.03)
Household size	8.44 (5.78)	8.66 (6.74)	-0.22 (0.43)
Number of children	4.37 (3.57)	4.40 (3.67)	-0.03 (0.26)
Is parent to some children in household	0.75 (0.43)	0.70 (0.46)	0.05* (0.03)
Has self-reported long-term health condition	0.35 (0.48)	0.40 (0.49)	-0.06* (0.04)
Urban	0.22 (0.41)	0.19 (0.39)	0.030 (0.03)
Annual consumption per capita (nominal USD, top 1% winsorized)	304.43 (329.75)	242.33 (247.38)	62.10*** (23.64)
N	1,601	210	1,820

** denotes significance at 10%, ** 5%, and *** 1%.*

In addition to the results of the understanding tests, our qualitative research did not reveal any common pitfalls that could lead respondents to misinterpret the question. This supports our conclusion that overall, this choice experiment was easier for respondents to understand than the VSL scenarios.

A small number of respondents only (this was raised by one individual in a single focus group in Migori, Kenya) may have failed to treat each subsequent choice as independent from the previous. Many respondents demonstrated reasonings consistent with thinking through the trade-off. In all focus groups, respondents demonstrated using ethical frameworks to make the trade-off, and considered and discussed whether to switch as the number of cash transfer changed. Finally, in the cases where people were unwilling to switch, people were able to support this with a clear reasoning suggesting that these data points are true preferences given the perceived benefits of saving a life and receiving cash transfers.

[128] The observations presented in Table 24 are all from respondents classified as “poor”.

Table 25. Qualitative data from focus groups exploring respondent understanding of cash vs life choice experiment.

Correct insight inferred*	Direct quotes from focus group discussions
Respondents were using ethical frameworks to make the trade-off between cash and life	*"I was picking the one that was saving one more life. This is because there is a big possibility that this child will support his/her family once they are well and grow up. It's not good for more families to be happy because they have money yet one family is mourning"*
Respondents considered switching when the number of cash transfers was changed	*"At some point I'll switch – if there's no money in the community, you can't even afford the hospitals and the vaccines; you need some money to live"*
Respondents supported non-switching behavior with moral reasonings	*"I was always picking the [life] even if you give cash transfers to the entire village [or] even to the entire Kenya with billions of people. If there is one sick child in Migori County that needs treatment, it's better to give all the money to save the child than give everyone in the county cash transfers. This is because that child can be a good person…"*

*We did not note any specific misunderstandings of this choice experiments from our focus group discussions.

ADDITIONAL ANALYSIS: SUBGROUP ANALYSIS

The table below includes subgroup analysis for this choice experiment. The point estimates for Ghana were substantially higher than for Kenya (as discussed in the main text, this reflects the higher proportion of non-switchers in Ghana). The estimates were also higher for poor households, and urban households.

Table 26. Subgroup analysis for cash vs. life choice experiment.

Variable[129]	Point estimate for the first 3 questions only (SE)
Central estimate vs	91.05 (44.19)
IPW estimate[130]	91.36 (-)
Ghana vs	200.88 (1352.40)
Kenya	14.50 (67.17)
Christian vs	86.76 (9.30)
Muslim	116.89 (15.79)
Urban vs	122.63 (36.56)
Rural	86.20 (6.26)
Richer	61.49 (5.20)
Poorer	111.97 (13.39)

[129] We omitted the parents/non-parents point estimates in because analysis for these variables wasn't possible (sample size for non-parents was too small).

[130] We estimated this value using inverse probability weighting to adjust for difference in demographics between our estimation subsample (those passing dominance and consistency tests) and the full sample. For this estimation, we did not apply GMM to obtain consistent standard errors.

ADDITIONAL ANALYSIS: SWITCHERS VS. NON-SWITCHERS

As noted in the main text, we found that the results of our estimation for this choice experiment are highly sensitive to the number of 'non-switchers' in our sample. Here we present the results of a deeper look at four of the questions that arise from this finding:

1. How do switching points vary across the full sample, compared to the Kenya only and Ghana only samples?
2. Is there any difference in demographics among switchers vs non-switchers?
3. How does our central estimate change if we focus on switchers only?
4. How does our central estimate change if we take a different analytical approach, that places more weight on non-switchers?
5. Is there any difference in VSL and the relative valuation of different age groups among switchers vs non-switchers?

How do switching points vary across the full sample, compared to the Kenya only and Ghana only samples?

Below we present the raw distribution of switching points among the estimation sample (i.e. those who pass the dominance and consistency tests). Due to the design of the choice experiment, we can only narrow down switching points to the interval within which switching occurred. For ease of comparison, the Ghana and Kenya graphs have the same scale for the horizontal axis.

Overall, we see that the majority of switchers switched between 1 to 10, 1 to 20, and 1 to 30. In fact, over 50% of switchers did so in these three groups below 30, or at an implied value of 30,000 USD. The distribution of switchers in Ghana is to the right of Kenya, meaning people typically switch at higher values.

Figure 11. Distribution of switching points across the full sample (N = 1,493)

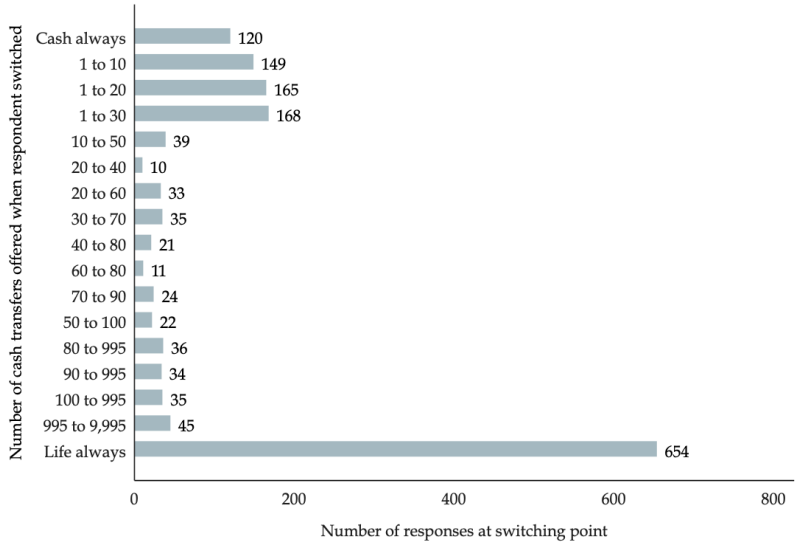
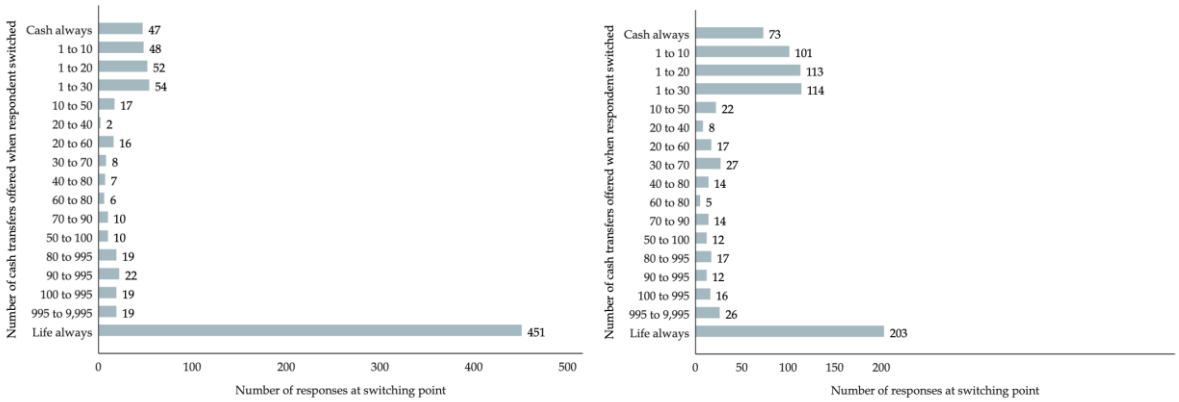


Figure 12. Distribution of switching points across respondents from Ghana (left; N = 790) and Kenya (right; N = 703)



Is there any difference in demographics among switchers vs non-switchers?

Below we compare the demographics of switchers and various types of non-switchers, from the subsample that passed the dominance and consistency tests. We find that "life always" non-switchers (who chose the program that saves an additional life no matter how many additional cash transfers are offered), are on average younger, more likely to be illiterate, and come from larger households with more children. We also find a higher proportion of this group are Muslim.[131] "Cash always" non-switchers are on average older, more likely to be illiterate, and more likely to have a long-term health condition.

[131] Note that each of these trends in demographics are also differences between our Ghana and Kenya sample. We haven't conducted any multivariate regressions to identify the strongest correlates associated with being a non-switcher.

Table 27. Demographics of switchers vs. "cash always" non-switchers and "life always" non-switchers.

Variable	Switchers	"Cash always" Non-switchers		"Life always" Non-switchers	
	Mean (SD)	Mean (SD)	Difference (SE)	Mean (SD)	Difference (SE)
Ghana	0.374 (0.484)	0.392 (0.490)	0.018 (0.047)	0.690 (0.463)	0.316*** (0.025)
Female	0.502 (0.500)	0.492 (0.502)	-0.01 (0.049)	0.563 (0.496)	0.061** (0.026)
Age	40.875 (14.887)	44.608 (15.888)	3.733** (1.467)	38.394 (14.391)	-2.481*** (0.768)
Can read	0.560 (0.497)	0.508 (0.502)	-0.052 (0.049)	0.384 (0.487)	-0.176*** (0.026)
Christian	0.614 (0.487)	0.642 (0.482)	0.027 (0.048)	0.583 (0.494)	-0.032 (0.026)
Muslim	0.276 (0.447)	0.258 (0.440)	-0.017 (0.044)	0.346 (0.476)	0.070*** (0.024)
Household size	7.823 (4.931)	7.350 (5.055)	-0.473 (0.483)	9.420 (6.695)	1.597*** (0.302)
Number of children	4.132 (3.279)	3.558 (2.961)	-0.573* (0.317)	4.830 (3.943)	0.698*** (0.188)
Is parent to some children in household	0.753 (0.431)	0.692 (0.464)	-0.062 (0.043)	0.754 (0.431)	0.000 (0.023)
Has self-reported long-term health condition	0.340 (0.474)	0.479 (0.502)	0.139*** (0.047)	0.329 (0.470)	-0.011 (0.025)
Urban	0.213 (0.410)	0.225 (0.419)	0.012 (0.040)	0.219 (0.414)	0.006 (0.022)
Annual consumption per capita (nominal USD, top 1% winsorized)	302.045 (308.553)	286.571 (218.325)	-15.474 (29.297)	310.712 (370.361)	8.667 (17.689)
Observations	827	120	947	654	1481

*, **, and *** indicate statistically significance at 10%, 5%, and 1% respectively.

How does our central estimate change if we focus on switchers only?

If we conduct the same logit estimation approach on our sample of switchers only, we find an average estimated value of -1.85 (as opposed to 91.05 for the full sample). The point estimate for this subgroup is negative with very large standard errors. We do not believe these reflect true preferences, but rather a result of fitting the logit model with sparse data in certain ranges. Hence, we recommend not focusing on these point estimates, but rather on the median values and distributions of switching points (presented in graphs above).[132]

Among switchers, we again see that the median is higher in Ghana than it is in Kenya.

[132] Note that due to overlapping intervals of values that respondents receive in their choices, it is sometimes difficult to pin down the medians to a narrower interval.

Table 28. Switching point estimates of the cash vs. life choice experiment by country

	Full sample		Kenya		Ghana	
	Mean (SE)	Median	Mean (SE)	Median	Mean (SE)	Median
Full sample	91.05 (44.19)	1 to 100	14.50 (67.17)	1 to 50	200.88 (1352.40)	>9995
Switchers only	-1.85 (31.70)	1 to 30	-10.10 (54.41)	1 to 30	17.60 (55.94)	1 to 50

The values for means, SE and medians represent the number of $1,000 cash transfers

How does our central estimate change if we take a different analytical approach?

In this choice experiment respondents faced a randomly selected group of choices, and the intervals created by these choices are often overlapping (e.g. one respondent may face pairs of choices that differ in the number of cash transfers by 1, 30, 70, 90, etc., and another respondent with 1, 50, 100, etc.). This makes it difficult to estimate an accurate median, as the true switching point may fall anywhere within a wide range.[133] Here we present a different approach to the analysis (that is different from typical practices seen in the literature) in which we make the assumption that an individual's switching point is at the mid-point of their switching interval (except at extremely high values where we take the lower bound).[134] We use these mid-points to estimate an approximate mean and median switching point across the sample.

The results when this approach is applied to the full sample, and to switchers only, is seen in Table 29. We find that the mean switching points across the full sample are extremely high, driven by large number of respondents who do not switch even when 9995 additional cash transfers are offered. When we restrict this estimation to the sample of respondents who were willing to switch between programs, we find a mean switching point of 84 cash transfers. This result is similar to that of the logit estimation (91) as both approaches diminish the effect of the skew towards high values.[135]

This alternate analytical approach again demonstrates the sensitivity of the results of this choice experiment to the extremely high values of the 38% of respondents who always prefer a life-saving programs.

[133] The median for the overall distribution, without taking the midpoints, is between 10 and 100 cash transfers. For Ghana it is above 10,000 cash transfers, and for Kenya is between 1 and 30 cash transfers.

[134] For example, if a respondent chose the program offering more cash transfers when the difference is 70, but switches to the program saving an extra life when the difference is 30, we assume their true value is 52.5. For those that do not switch to the program saving an extra life when there is a difference of 1, we assume a value of 0. For those that do not switch to the program offering more transfers above 80 we assume the lower bound rather than mid-point (since these values are very high, and have a very large impact on the results).

[135] Here, we drop the values of non-switchers, while in the logit model we only include data from the first 3 choices, so non-switchers values are treated as above 100, but not given a specific high value.

Table 29. Choice experiment results taking an alternate analytical approach, based on midpoints rather than logit estimation.

		Full sample	Kenya only	Ghana only
Full sample	Mean switching point (number cash transfers)	4127	3124	5623
	Implied value using mean (USD)	$4,126,550	$3,124,670	$5,622,920
	Median switching point (number cash transfers)	~80	~30	>9995
	Implied value using median (USD)	~$80,000	~$30,000	>$9,995,000
Switchers only	Mean switching point (number cash transfers)	84	76	97
	Implied value using mean (USD)	$84,487	$76,349	$96,932
	Median switching point (number cash transfers)	~14	~12	~17
	Implied value using median (USD)	~$14,000	~$12,000	~$17,000

Is there any difference in VSL and the relative valuation of different age groups among switchers vs non-switchers?

Finally, we examine difference in the results of our other primary methods according to the response given to this choice experiment.

We found that:

1. In their responses to the VSL, "life always" non-switchers do not express higher values than switchers, except for children 5 or older.[136]
2. In their responses to the relative values of age groups, "cash always" non-switchers exhibit the highest relative values of individuals aged 5 to 18 and individuals aged 19 to 40 relative to those under 5, compared to all other groups. In contrast, "life always" non-switchers exhibit the lowest relative values of these individuals (relative to individuals under 5).
    a) These results suggest that "cash always" non-switchers place lower relative values and absolute monetary values on the lives of individuals under 5 years.
    b) In contrast, "life always" non-switchers place higher relative values and absolute monetary values on the lives of individuals under 5 years.

[136] We are only describing the mean values by group here; we have not done statistical tests for the differences between groups as there is a large number of combinations.

Table 30. Comparing VSL and Relative values of age groups across different groups of response types for the choice experiment.

	Full sample for "cash vs. life" choice experiment	Switchers	"Cash always" non-switchers	"Life always" non-switchers
N	1601	688	120	654
I. Comparison of VSL and money vs life choice experiment				
VSL self (5/1000, model A)				
Mean (USD)	34867.92	36046.66	38977.24	29568.87
SE (USD)	2482.96	3415.19	8263.20	3688.97
N	498	223	31	210
VSL child (5/1000, model A)				
Mean (USD)	36477.85	38373.58	30636.06	34038.58
SE (USD)	3644.97	6344.75	9071.44	4994.86
N	435	196	27	178
VSL child under 5 (5/1000, model A)				
Mean (USD)	40789.42	46957.28	37238.2	30689.87
SE (USD)	6608.97	13892.83	18304.66	5695.96
N	204	85	12	93
VSL child 5 or older (5/1000, model A)				
Mean (USD)	32670.22	31800.47	25354.34	37702.47
SE (USD)	3609.70	3498.25	7745.58	8420.47
N	231	111	15	85
II. **Comparison of VSL and relative value choice experiment**				
Relative values of age groups (relative to under 5)				
N for relative values	1344	688	94	562
Value of 5 to 18	0.78	0.83	1.02	0.64
SE	0.04	0.05	0.13	0.10
Value of 19 to 40	0.10	0.24	0.60	-0.27
SE	0.05	0.05	0.15	0.16
Value of above 40	-0.71	-0.40	-0.23	-1.49
SE	0.12	0.11	0.22	0.44

APPENDIX 4: COMMUNITY PERSPECTIVE: RELATIVE VALUE OF AGE GROUPS

In this appendix we provide an in-depth description of the approach we took to capture the relative value of age groups (with our community perspective choice experiment), provide more data related to method credibility, and present the results of additional data analysis.

METHOD OVERVIEW

First, respondents were introduced to the scenario. We specified:

1. That decision makers often have to make difficult trade-offs between different kinds of life-saving programs, and provide a number of examples.
2. That life-saving programs may target people in different age groups.
3. They should assume that they are the village chairman and need to choose between to available life-saving programs for their community.
4. The cost of both life-saving programs is the same.

Next, respondents were asked to make a choice where one option is clearly dominant, in order to assess basic understanding of the scenario:

"Program A saves 200 lives of people aged 19-40 years; Program B saves 100 lives of people aged 19-40 years. Which one would you choose?"

Respondents were then asked to make 6 similar choices, in which the age group and the number of lives saved changed for each question (the varying attributes are highlighted below in blue). Each respondent was randomly assigned to one of 8 choice sets. So, across the sample 48 different comparisons were made. After the first two choices we asked respondents to explain the reasoning behind the decision. A visual aid was used throughout to demonstrate the choice to the respondent.

"Program A saves [100/200/300/400/500] lives of people aged [under 5/5-18/19-40/over 40], Program B saves [100/200/300/400/500] lives of people aged [under 5/5-18/19-40/over 40]. Which one would you choose?"

Finally, the respondent made a final choice that was a variation of a previous choice, in order to check for internal consistency in their response. For example, if in an earlier choice a respondent chose a program that saves 100 lives of children under 5, over a program that saves 400 people aged over 40, they were asked:

"Program A saves 100 lives of people aged under 5 years; Program B saves 200 lives of people aged over 40 years. Which one would you choose?"

If the respondent switched to the program saving older lives, which contradicted their earlier choice, we classified this respondent as inconsistent.

METHOD VALIDITY

Similar to the monetary value of life choice experiment, we found good evidence of respondent understanding. Consistency of preferences was not quite as high as the first choice experiment, but we still found that a large proportion of our sample was consistent across the 6 choices presented

(84.8%). Overall, data from 82.3% of respondents was included in the final estimation, and we found no statistically significant differences between those included and those excluded.

Table 31. Summary of validity test results for relative lives choice experiment.

Test	Result
Dominance Test	97.1% of respondents picked the correct choice when presented with a dominance test choice.
Consistency Test	84.8% of respondents were consistent across the different choices presented to them.
Both tests	82.3% respondents passed both tests.
Sample balance test	There were few statistically significant demographic differences between the sub-samples that passed the dominance and consistency tests, and the full sample (see Table 32).

Table 32. Comparing respondent characteristics between the estimation sample and the rest of the sample.[137]

Variable	Relative Values Estimation Sample (SD)	Rest of Sample (SD)	Difference (SE)
Female	0.53 (0.50)	0.58(0.49)	-0.06* (0.03)
Age	40.09 (14.87)	41.641 (15.50)	-1.55* (0.92)
Can read	0.46 (0.50)	0.56 (0.50)	-0.10*** (0.03)
Christian	0.60 (0.49)	0.66 (0.48)	-0.06** (0.03)
Muslim	0.29 (0.46)	0.30 (0.46)	-0.01 (0.03)
Household size	8.50 (5.74)	8.30 (6.61)	0.20 (0.36)
Number of children	4.38 (3.49)	4.36 (3.97)	0.02 (0.22)
Is parent to some children in household	0.74 (0.44)	0.74 (0.44)	0.01 (0.03)
Has self-reported long-term health condition	0.34 (0.47)	0.41 (0.49)	-0.068** (0.03)
Urban	0.21 (0.41)	0.21 (0.41)	-0.00 (0.03)
Annual consumption per capita (nominal USD, top 1% winsorized)	290.81 (299.01)	326.06 (410.71)	-35.250* (19.83)
N	1,493	320	1,820

* denotes significance at 10%, ** 5%, and *** 1% levels

Our qualitative research revealed a couple of relatively any common pitfalls that could lead respondents to misinterpret the question. Some respondents failed to treat each subsequent choice as independent from the previous and others might have felt like the choice sets were too many.[138] However, this was only expressed by a few respondents and many others demonstrated reasonings consistent with thinking through the trade-off. In all focus groups, respondents demonstrated using ethical frameworks to make the trade-off, and considered and discussed which age groups were most

[137] The observations presented in Table 32 are all from respondents classified as "poor".
[138] Once this issue was identified, we asked enumerators to stress the independence of each scenario, which we think minimized any impact on our results.

preferential. Evidence of good and poor understanding from focus group discussions is presented below, see Table 33.

Table 33. Qualitative data from focus groups, exploring respondent understanding of the relative value of different ages choice experiment.

I. **Evidence of good understanding**	
Correct insight inferred	Direct quotes from focus group discussions
Respondents were comparing both quantity and age of people in each choice set presented to them	*"I was looking at both the quantities and the ages. But I was mostly picking the children because they are the future generation"*
Respondents provided detailed reasons for favoring one age group over the others	"*The ones who are under 5 will die if their dependents die anyway. At least the 5-18 year olds can take care of themselves without necessarily having other people around"*
Respondents supported non-switching behavior with moral reasonings	*"I'll never switch, because the adults can fend for themselves. Not even 1,000 I won't prioritize the adults"*
Respondents supported extreme choices with moral reasonings	*"Even if you put 1 billion adults, I will still pick children because they are the future generation and the joy of their families"*
II. Evidence of misunderstanding	
Type of misunderstanding	Direct quotes from focus group discussions
Respondents failed to treat each new choice as independent from the previous	*"Saved different ones each time i.e. if I saved young ones before I will save old ones." "But at the end I picked the adults because I thought I had saved children too long."*
Survey fatigue	*"The choices were too many; it got to a point where I just answered to finish"*

ADDITIONAL ANALYSIS

The table below includes subgroup analysis for this choice experiment. Of note, non-parents and urban respondents, place greater weight on older age groups than individuals under 5, than parents and rural respondents.

Table 34. Subgroup analysis for relative value of different age groups choice experiment

Variable	5 to 18, relative to 0 to 5 (SE)	19 to 40, relative to 0 to 5 (SE)	Above 40, relative to 0 to 5 (SE)
Aggregate vs IPW[139]	0.791 (0.002) 0.795 (-)	0.113 (0.002) 0.124 (-)	-0.712 (0.014) 0.670 (-)
Ghana vs Kenya	0.636 (0.08) 0.898 (0.002)	-0.027 (0.010) 0.200 (0.002)	-1.523 (0.154) -0.267 (0.007)
Christian vs Muslim	0.845 (0.002) 0.722 (0.011)	0.131 (0.003) 0.126 (0.011)	-0.588 (0.016) -1.093 (0.141)
Urban vs Rural	0.914 (0.005) 0.747 (0.003)	0.273 (0.004) 0.058 (0.003)	-0.449 (0.030) -0.801 (0.022)
Parents vs Nonparents	0.743 (0.003) 0.916 (0.005)	0.105 (0.003) 0.134 (0.006)	-0.757 (0.022) -0.589 (0.037)
Richer vs Poorer	0.821 (0.001) 0.840 (0.003)	0.162 (0.001) 0.163 (0.003)	-0.629 (0.009) -0.675 (0.021)

[139] We estimated these values using inverse probability weighting to adjust for difference in demographics between our estimation subsample (those passing dominance and consistency tests) and the full sample. For this estimation, we did not apply GMM to obtain consistent standard errors.

APPENDIX 5: MORAL REASONING DATA

Below we summarize the range and prevalence of justifications used by beneficiaries when making moral trade-offs. The intensity of the color in the Full sample, Kenya and Ghana columns, indicates the approximate prevalence of each justification (darker = this type of justification arose more frequently). We have also made a tentative attempt to group together these justifications into broader decision frameworks. We have bolded the justifications and related framework that appeared most often, and so were described in the main text of Section 3.

Note: We are aware there is a lot of subjectivity in the coding and grouping of these justifications. As such, we verified our coding by having a second independent reviewer code a random sample of 25% of the qualitative interviews. The same patterns arose, although there were some minor semantic differences in definitions particularly of the 'frameworks' used. As such, where discussed in Section 3, we place more emphasis on the specific justifications (which relate directly to the reasons given by the respondents) than the specific frameworks (which are grouped and defined by us as researchers as opposed to the respondents).

Table 35. Breakdown of justifications and prevalence of those justifications, where respondents made comparisons between life-saving and consumption-increasing interventions.

	N	Justification	Full sample	Kenya	Ghana	Quote	Framework
In favor of consumption-increasing interventions	1	Families will use the cash productively to improve their outcomes				*I would definitely choose the cash. No matter what, 80% of those beneficiaries will be good! Nobody wants to get help and waste it and come back to the same route; you will not do that; everyone needs a push. The cash transfer is a push, it's motivation.*	Present-value orientation
	2	**Present value of cash outweighs future potential value of saved life**				*The cash makes a household better off. This is because they have built a good house, they now eat good food, they have money and food. If the child is saved, the family is still at the bottom of the ladder because they still have very little money.* *You don't know whether the child is coming to save you; if you miss that, you could waste a lot; so he could be a savior; it could be the next Rawlings [ex-president of Ghana]; But you still always take the cash so you are not in poverty; because you don't actually know if that child will be Rawling*	

	3	Cash is a sustainable intervention that allows for renewable investments				*In general, the government should prioritize programs that improve the living conditions of people over the ones that reduce the risks of dying. I have learnt that empowering people to get an income on their own is very sustainable. If communities were empowered, they would not need NGOs to keep supporting them because they would have their own income. Additionally, when people have their own income, they can even buy their own ITNs and deworming pills to reduce their risks of dying.*	Outcome -based orientation
	4	Cash directly alleviates poverty				*Cash improves the living standards of the poor household(s) in the community. Money will help them out of poverty. They are ready to work, but don't have enough work.* *Because money is used in education, feeding, and buying farm produce. So, when more cash transfer is given, more people benefit.*	
	5	Cash can be shared beyond a family, while an ITN benefits only one individual				*The one household that will receive the cash can help so many people in the village, so many more than the life of one child.* *Many people are poor and once you help them, they will also help the children and other less fortunate people in the community.*	**Greatest-good orientation**
	6	**Better to help many in the family than to only save one in the family**				*If the one house is supported more people are supported as a household comprises of many, like mother, father, and several children* *Cash given to one family impacts more than one life. It serves a family and not an individual.*	
	7	Cash more versatile (i.e. business, school fees)				*More cash transfers that will help them buy food, build a good house, buy clothes and pay fees.* *More cash transfer will help solve many more problems/challenges which may include saving at least one more life.*	
	8	Cash amount ($1000) is more than enough to make the poorest better off (can actually give less and spread across community)				*[The cash] can do a lot! It can buy foodstuffs, it can be used to build households; take care of house; business; this is more impactful than saving one life; the poorest households don't need that much to be better off – I would give less than 5,000 GHS to each, and spread it further – even if I give 1,000 cedis it would support.*	Outcome orientation
	9	Fate will select a child's death; death inevitable for everyone				*No matter how it is, people will die so I prefer we enjoy ourselves before we die. Death is inevitable, when there is money people can access good health care; God will choose the saviors; it's fate; He will ensure the promising people will not be the ones to catch the disease.*	Religious orientation
	10	Birth is common; cash is rare				*Children can always be born to fill the earth again if they die; we now need money to survive, the lifestyle here is getting difficult.*	Needs orientation

	11	Malaria nets are not guaranteed to save life immediately.				*Sometimes you have the bednets but you still get bitten and infected.* *But the bednets have different types – the round one is not good, not quality, I don't trust bed nets.*	Skeptical of nets
Neutral	12	Allocation decisions should be made based on the present needs of the community				*All these programs are good and beneficial, so I wouldn't pick one: I would distribute the money according to the needs of every family.*	Needs orientation
In favor of life-saving interventions	13	**Children have future potential economic value**				*This is because children are very important—they are the leaders and economic forces of the future. The one child you save could be the one God blesses to provide and take care of the family in future.* *The child saved can become President and can help in developing the community.*	**Future-value orientation**
	14	Children have real economic value to a family unit				*When they are a little grown and can support in house chores and other minor works in the community.*	
	15	Satisfaction derived from life translates to economic good				*Even if you give them everything they need and the child is sick, they will not enjoy.*	
	16	Death is an economic burden				*Save the life of the child because even the funeral expenses is likely to use up that cash support*	
	17	Death is a severe emotional burden that disrupts the family				*Because it's quite painful and emotional as a parent to watch a child die, and this will always trump anything else (improved lives etc.).; Because I am a father if my child dies or is sick it affects most of my work.*	
	18	**Deontological rule that life is priceless**				*It's worse for the child to die than for the family to keep suffering—life is more important than anything else.*	**Deontological view to life**
	19	**Sanctity of life; sacrificing life is satanic; regards having children as the purpose in life**				*Money is satanic and life is more important. You cannot compare life with anything.* *When you have life, God been faithful, that life will grow to be the assist the community, but if one has all these cash and doesn't have any extra life or child to spend it what is the profit in having all these cash?*	
	20	**Cash may be misused**				*Most poor families will misuse cash transfers if given directly-- they will use it to buy things they did not have. They will start business they do not have any experience in. They will buy the things they do not have but these will not help them.*	**Skeptical of cash improvements**
	21	Cash can create community tensions				*The money is not enough for everyone, in our part of the world, people get aggravated if they are not beneficiaries, and it will cause a lot of conflicts in the community, even though there is wealth variation.*	

	22	Cash is non-sustainable/non-renewable				*Today, if I were to receive a program, I would pick the bednets because I don't feel that I will ever have enough money – 5,000 GHS ($1000) would still leave me wanting. Cash transfer has more risks than you can imagine; I can invest that in farming, but what if the yield is bad, then I have thrown the money away, wasted it.*	
	23	Cash transfers of $1000 are not enough to reduce poverty				*Children require more help than just money to be used in one season. Even the amount of money they receive in cash transfers will not be enough to fully support the children anyway, so I would rather children's lives are saved.*	
	24	Poverty is the status quo and people survive				*Poverty is a problem but people die from things like malaria, not necessarily poverty. Poor people have been poor for a while and they have still been surviving.*	
	25	Death is always worse than poverty (cannot come back from death)				*Money can be sought but death is irreversible.*	Tangibility orientation
	26	ITNs have tangible, direct benefits that cannot be wasted				*I feel more pressure when I receive the cash transfer, my discretion for investment could be poor. Bednets are a guaranteed benefit.*	

Table 36. Breakdown of justifications and prevalence of those justifications, where respondents made comparisons between interventions saving lives of different ages.

	N	***Justification***	***Prevalence***	***Quote***	***Framework***
In favor of child-saving interventions	1	Children are innocent		*They are children and they are innocent. They have no sin.*	Deontological view to children's life
	2	Unfair to not give child full life while adult already enjoyed		*I have lived and can die but my daughter still has more years to live.*	
	3	Natural succession is for children to survive and replace adults		*"Child is more valuable because of replacement; we believe that when you take care of the child, he replaces you in your old age, at present people marry because of child; if there is a problem and no child, you see that your effort is wasted, what else is there."*	
	4	Parental moral responsibility to save child		*My children are more important to me than myself*	

	5	Future potential value of child is greater than adult		*"I want to spend all of the money on the bednets because it will secure our future, because the children are our future; so that they can be healthy, go to school, and save the entire community"*	Economic view to child life
	6	Child is an asset to the family because will care for adults		*They will grow to help the elderly. They will dig my grave when I die.*	
	7	Adult potential already realized and exhausted		*The kids need to grow to also experience life but the aged are already grown and must pave way for the new and younger generation.* *We struggle for those that are behind us, not in front us; Your father educated you, and so on*	
	8	Need to protect the weakest individual in the family		*Children are unable to articulate their needs and hence the need to seek out their interest* *Children are prone to many diseases* *Children between 5-18 deserve to be helped as are not as strong yet*	Weakest link view
	9	Emotional burden of death of a child is greater than adult		*It's more painful to a parent to lose a baby than to lose a grown up*	Emotional view
Neutral	10	Accounts for real constraints that make it more/less valuable to have children/adults saved in a given community		*Most of children's medicine are cheaper so I will buy more of that one.*	Pragmatic view
In favor of adult-saving interventions	11	Present economic value of adult caretaker is greater than child (greater asset to the household)		*They are mostly providers of the households* *They will take care of the lives of those who remains*	Present conditions and outcomes view
	12	Already invested in these individuals to achieve their potential		*We have suffered for this age group to get to this level so I will support them. The other are still young and much hasn't been invested there yet*	
	13	Adolescents and adults are easier to care for		*It's difficult to maintain under 5 kids.*	
	14	That these individuals have survived past 5 means they are strong		*We have suffered for this age group to get to this level so I will support them. The other are still young and much hasn't been invested there yet*	

	15	Older children can take care of those below and above them		*The older children have a chance of helping other people or assisting than the younger ones*	
	16	Adults are more prone to disease and need assistance		*The 5-18 are more vulnerable and are at high risk to disease or accident than the under 5 years*	
	17	Elderly need as much support as possible		*They are old and frail and so need help/support to survive*	
	18	Moral duty to care for your own parents		*They gave birth to me, I need to care for them too.*	Duty to family
	19	I am in this age bracket and I want to save myself		*I am that age.*	Self-interested view
	20	You can have more children		*Can still give birth to more children* *If I save the 19 to 40, then I'll be able to give birth to kids that will be under 5 again. If you safe the under 5, they will need not less than 10 years to grow and give birth.*	

APPENDIX 6: SUBJECTIVE WELLBEING- ADDITIONAL DATA

In this appendix we provide additional context for Section 5: Subjective Wellbeing. Table 37 summarizes the main reasons respondents gave for their self-reported life satisfaction ladder scores. Most respondents associate their life satisfaction ladder score with the availability of economic assets. For example, the most common reasons for low self-reported life satisfaction scores are lack of money, lack of businesses/jobs, lack of proper shelter and lack of food.

Table 37. Reasons given for self-reported life satisfaction ladder scores

Reason given	Freq. mentioned (N=1751)	% of mentions
Lack of money	1321	75%
Problems with job/business (including lack thereof)	671	38%
Lack of education for themselves or their children	490	28%
Problems with their house (including lack of water/electricity/furniture etc.)	548	31%
Lack of food	559	32%
Poor health (including concerns about accessing care)	271	15%
Stress/Emotion (including lack of peace of mind/the strain of finding money)	110	6%
Dependency on others (i.e. lack of independence)	91	5%
Lack of a spouse (i.e. husband, wife, boyfriend, girlfriend)	131	7%
Lack of land for farming	174	10%
Lack of livestock e.g. cows and goats	112	6%
Lack of poultry e.g. chicken	29	2%
Lack of religion/God	5	0.3%
Lack of a car/motorcycle	21	1%

ADDITIONAL HYPOTHETICAL SCENARIOS

In addition to observing the effect of respondent characteristics to life satisfaction scores, we captured subjective views of respondents to different hypothetical scenarios. Similar to current conditions, we observed associations between these hypothetical conditions and respondents' life satisfaction scores.

Increasing Consumption In order to capture the effect of hypothetically increasing future consumption on self-reported life satisfaction, we asked the following question:

> *You said that you were on step X of the ladder. Now imagine a family that is very similar to yours (same size, in the same community, similar income etc.). However, the difference is that one year ago they received a cash transfer for $1,000. With this money they were able to invest in a small business so their regular income has increased a little (roughly double what yours is today), and*

they completed some basic repairs to their house. Where on the ladder would you say the members of this family stand today?

On average, respondents noted that a $1000 cash transfer would move a hypothetical family up the ladder by 2.9 steps.[140] Across respondents in Kenya and Ghana (n=170), the average expected life satisfaction after receiving a cash transfer of $1000 is 5.91 compared to the current average life satisfaction of 2.97. Since economic factors (i.e. money, businesses etc.) primarily influence life satisfaction (for about 71% of respondents), it would mean that for these respondents, receiving a cash transfer of $1000 would hypothetically double their life satisfaction.

Averting death In order to capture the effect of hypothetically averting the death of a child on life satisfaction, we asked the following question:

You said that you were on step X of the ladder. Now, again imagine a family that is very similar to yours (same size, in the same community, similar income etc.). They have NOT received any cash transfer, and their income is the same as yours is today. However, one year ago, one of the children was very sick and would have otherwise died, but they received support from a health charity that provided the child with the required treatment and the child is now healthy. Other than the support to that child, the family is unchanged. Where on the ladder would you say the members of this family stand today?

On average, respondents noted that saving the life of a child under 5 who would otherwise die moves a hypothetical family up the ladder by 1.2 steps. Across respondents in Kenya and Ghana (n=170), the average expected life satisfaction after averting the death of a household child under 5 $1000 is 4.19 (Range: 0 to 10) compared to the current average life satisfaction of 2.97.[141] We do not put much weight on these values as it is difficult for us to describe and for respondents to conceptualize the impact of avoiding a bad outcome. A more robust way to estimate the impact would be to look for associations between life satisfaction scores and recent deaths in a large dataset of relevant respondents, which was beyond the scope of this study.

[140] Note that while a cash transfer of 1,000 USD roughly doubles consumption for one person for one year, this perceived increase in life satisfaction ladder is significantly higher than the estimated coefficient of doubling consumption (which is 0.42) based on a regression, though the latter is not a causal estimate.

[141] However, this question is ambiguous so the numbers cannot be interpreted directly. We struggled to find a less ambiguous framing that captures the same idea. We expect that greater value comes in the comparison of the qualitative across the two scenarios. We plan to analysis this data and include in the final version of the report.

LIFE WORSE THAN DEATH

In order to assess whether respondents perceive of a life satisfaction score that is so low that life is worse than death, we asked a sub-sample of respondents (n = 70) the following question:

> *Sometimes people's lives are very difficult, and they struggle every day – is it ever possible for someone to lives such a difficult life that it is worse than dying? If yes, at what point on the ladder from 0 to 10 is it worse than dying? What are the characteristics of this life?*[142]

Approximately one third of respondents stated that it's not possible to have a life that's worse than death. These respondents cited deontological frameworks such as the inherent and immeasurable value of life regardless of other factors.

The remaining respondents (close to two thirds) indicate that there are points on the ladder where life is worse than death. For these respondents, this point is substantially lower than their current life satisfaction scores –the average point identified was 0.56 on a ladder from 0 to 10[143], compared to their current average life satisfaction score of 2.21.[144]

For many respondents, the point on the ladder that is worse than death is mainly characterized by lack of basic necessities (such as lack of food, shelter, clothing etc.). This closely mirrors the overall finding that subjective wellbeing is associated with living conditions. However, there are other themes that feature frequently in respondents' qualitative definers of the life worse than death. These include dependency on others, inability to provide for dependents, general suffering, mental anguish and poor health.

COMPARISON OF OUR RESULTS WITH OTHER STUDIES

The life satisfaction results for our study are substantially lower than expected, based on extrapolation from the results of nationally representative surveys in Kenya and Ghana. As shown in Table 38, the observed difference in life satisfaction scores between our results and national results is far higher than expected based on the observed relationship between income and life satisfaction.

We do not have a clear understanding why this is the case based on the data from this study alone. Two potential hypotheses include:

1. As this question was asked at the end of the survey for all respondents, the preceding questions either:
    a. Made respondents think about things that are lacking in their lives (e.g. through discussion of the impact of cash transfers), biasing their results down,
    b. Made respondents think that they may be more likely to receive discussed interventions if enumerators think their quality of life is particularly low.

[142] Enumerators were trained to only ask this question if respondents appeared at ease during the interview. We restricted it to our qualitative interviews only, so this sensitive topic was only broached by our strongest enumerators.
[143] This is an average value from two sub-samples. In the first sub-sample (N = 22), the average value is 0.69 and in the second sub-sample (N = 38), the average value is 0.42.
[144] There were 9 respondents (13%) who see their lives at or below the point worse than death.

2. The relationship between income and life satisfaction is not linear. Life satisfaction may drop more drastically as income levels reach as low as seen in the sample of this study.[145]

Table 38. Comparison of IDinsight study life satisfaction results to other studies.

	Reference figures		Observed difference	Expected difference based on regression coefficient on log annual consumption per capita[146]		
	IDinsight 2019	**National survey[147]**	**National – IDinsight**	**IDinsight 2019 (0.6)[148]**	**Deaton 2008 (0.75)[149]**	**Stevenson 2013 (0.25)[150]**
Kenya	2.29	4.66	2.37	1.04	1.31	0.44
Ghana	3.30	5.48	2.18	1.20	1.50	0.50

[145] It has been beyond the scope of this work to further explore whether this is the case within the datasets in other country level studies of life satisfaction.

[146] The equation we used to calculate the expected differences is: regression coefficient * ln (1710.5/300) for Kenya and regression coefficient * ln (2202.30/300) for Ghana. In the equations, 1,710.50 represents Kenya's GDP per capita (current, or nominal, USD), whereas 2,202.30 represents Ghana's (https://data.worldbank.org/indicator/NY.GDP.PCAP.CD). Across both countries, our respondents have approximately 300 USD (nominal) per capita annual consumption.

[147] World Happiness Report 2019, https://ourworldindata.org/grapher/happiness-cantril-ladder

[148] See Section 4 for the regression coefficient on log annual consumption per capita (0.6). This regression includes respondents classified as both "poor" and "wealthy" in order to estimate this relationship, though the main sample (and the one for which life satisfaction values are reported) is from "poor" respondents.

[149] The estimated coefficient on log real GDP per capita is 0.75, higher than our estimate. See Table 2, column 5 of https://pubs.aeaweb.org/doi/pdfplus/10.1257/jep.22.2.53.

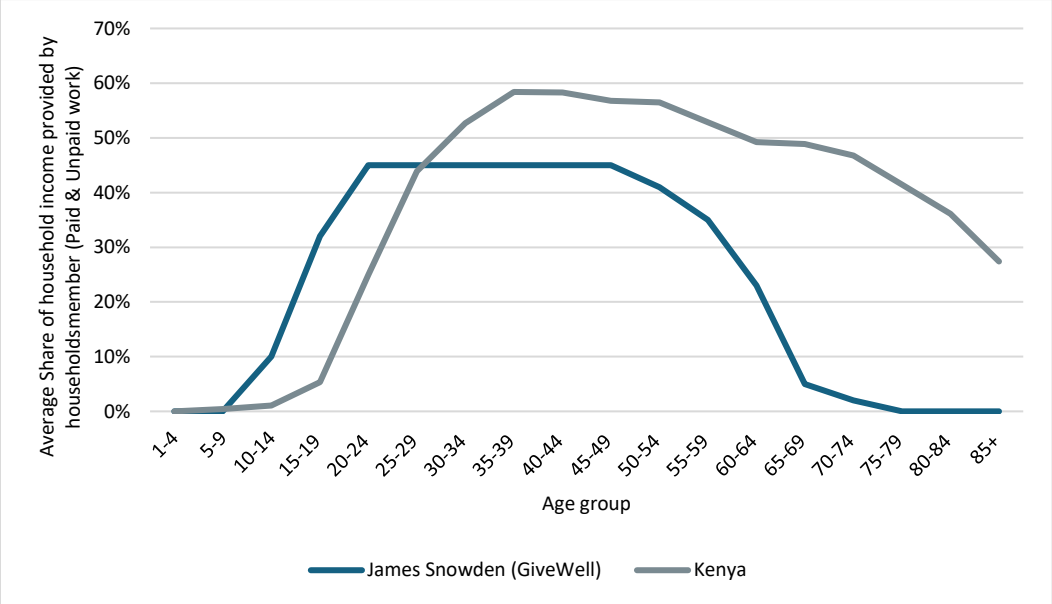
[150] For a cross-country regression on lower income countries, their estimated coefficient is 0.25 on log GDP per capita (PPP). Check Figure 1, upper panel, of https://www.nber.org/papers/w18992.pdf.

APPENDIX 7: ECONOMIC AND EMOTIONAL IMPACT OF DEATH

ECONOMIC IMPACT SECONDARY DATA ANALYSIS: STEP-BY-STEP RESULTS

First, we calculate the economic contribution of each household member by adding wages from paid activities and the imputed remuneration for unpaid work (e.g. house work).[151] The share of economic contribution in total household income is plotted by age group in Figure 13 alongside Snowden's model of economic contribution.[152]

Figure 13. Average economic contribution of household members as a share of total household estimated income, by age groups.



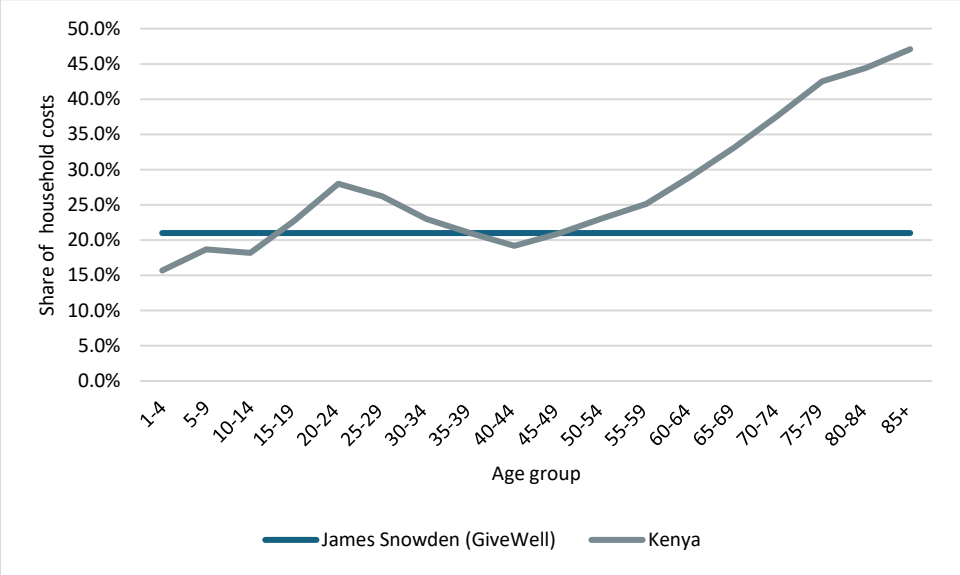
Compared to the Snowden model, the age - economic contribution relationship for poor households in Kenya is shifted towards the right. This means that, in Kenya, economic contributions to household income are on average lower for individuals aged 30 or less, and higher for those aged over 60.

[151] Unpaid work wages were estimated using time-use data on household activities, and average casual wage labor rates (adjusting for the age of the worker). Hourly wage rate was adjusted according to the age of the individual, according to recommendations in the literature: based on the recommendation by Whittington and Cook (2019). We assume that children under 5 do not contribute to any paid activities, and the value of time for unpaid work completed by children aged 5-17 is estimated at maximum 50% of that assigned to a male adult. The values add up to more than 100% across age groups, as for each age group we took the average among households with any member from that group. An alternative would be to pool across all 'poor' households in the data to account for distribution of age in the population. We choose to follow Snowden's approach for comparability and because the latter approach is not necessarily better as it fails to capture the fact that in some households due to the absence of certain age groups other age groups have a more significant income contribution.

[152] Snowden's model of economic contribution by age group can be found on row 45 of https://docs.google.com/spreadsheets/d/1FWqr1sp5asEwLQVpgicT2us-HZv44AuOEzLO86ZHXYE/edit#gid=1266107896.

We then calculate the costs for household members, including education, health and food expenses.[153,154] The share of cost in total household consumption is plotted by age group, alongside Snowden's model of costs, in Figure 14.[155]

Figure 14. Average costs of household members as a share of total household costs, by age groups



Next, we calculate the net current economic contribution of household members, plotted by age group,[156] alongside Snowden's model, [157] in Figure 15.

[153] Health costs aggregates in- and out-patient expenses. Food expenditure uses adult equivalent food consumption for each household adjusted for age and gender. Berti (2012) finds food consumption of children under the age of 5, and children aged 5-17 corresponds to 0.74 times and 1.04 times that of an adult man, respectively. For adult women, food consumption is estimated at 0.87 (Berti, 2012). The method here is the same for economic contribution, namely for each age group, averaging across households that have someone in the age group.

[154] The Kenya dataset does not contain data on other sources of costs (e.g. clothes, transport). It has a module on domestic tourism, but as ~98% of individuals had not carried any trip within the past 3 months, we decided not to include this source of costs.

[155] Snowden's model of cost (consumption) by age group can be found on row 46 of the same spreadsheet as economic contribution. It assumes constant consumption regardless of age.

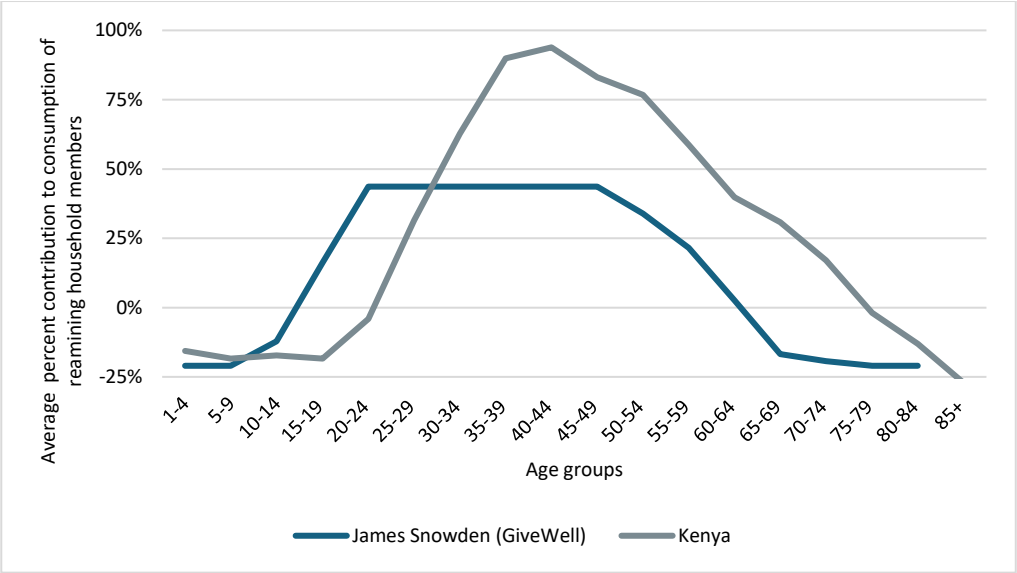
[156] We followed Snowden's calculation, using the following formula:

$$y = \left[(1 + x) \times \left(\frac{s - 1}{s}\right)\right] - 1$$

Where y is the increase (decrease) in consumption for other household members, x is the percent economic contribution (cost) to the household income, and s is the household size.

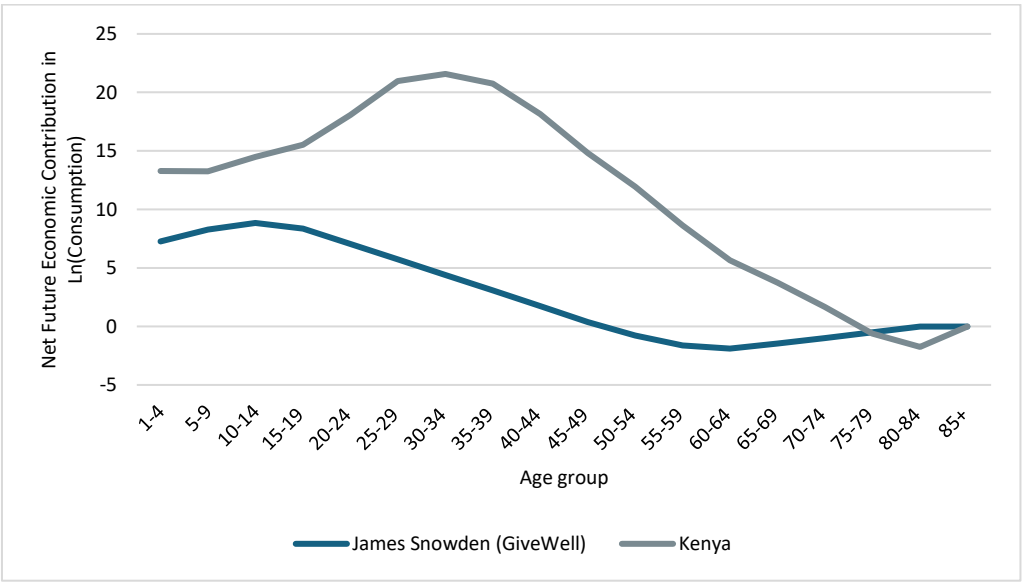
[157] Row 47 of the Google Sheet referenced above.

Figure 15. Average percent net contribution to consumption of other household members, by age group



Finally, we calculate the net future economic contribution of household members in terms of ln(consumption), accounting for the number of years spent in each age group and the survival rates of each age group. [158] This is plotted by age group below, alongside Snowden's model. [159] We differ from Snowden's model here by including the person's 'present' economic contribution during the second half of the age bracket, assuming for simplicity that any death during this period occurs in the middle in expectation.

Figure 16. Average net future economic contribution to the household by household member of different ages.



[158] We assumed a discount rate of 4.2% corresponding to the median value of staff inputs for discount rate in the GiveWell CE model.
[159] Row 49 of the Google Sheet referenced above.

The distribution of net future economic contribution of household members in terms of ln(consumption) is above Snowden's model for all age groups until age 75, and shifted to the right.

ECONOMIC IMPACT PRIMARY DATA ANLAYSIS: QUALITATIVE HEAT MAP

Below we summarize the themes that emerged from our qualitative analysis of the economic impact of death of individuals on household members. The table collates (a) all the themes that emerge along with short descriptions, (b) key quotes to help illustrate those themes, and (c) a color scale that indicates the overall prevalence of each of the themes across the 47 qualitative interviews. The selected quotes were randomly selected from some of the most descriptive illustrations of the themes across all the interviews we conducted. The intensity of the color indicates the approximate prevalence of each justification. The darkest color indicates that the theme arose more prevalently (it was mentioned by more than half of the respondents). The lightest color indicates that the theme arose less prevalently (it was mentioned by less than an eighth of the respondents).

Among the themes that emerged from our qualitative analysis, we noted a distinction between short- and long-term economic effects of death. Short-term effects are associated with the immediate and direct loss of income and expenses associated with the burial ceremonies of the decease. Long-term effects involve the future socio-economic implications of death on the household and community. Of note, the economic effects of death are not only negative; there are a few positive effects, marked with an asterisk.

Table 39. Breakdown of the most common themes that emerge from qualitative data on the economic burden of death.

	N	Theme (bold) and description	Quote	Prevalence
Short-term economic effects	1	**Selling or renting household assets** such as land, livestock, trees and farm produce in order to get enough money to cater for funeral costs.	*I cut down all trees around his house and sold them for Ksh.9,000* *[The deceased] used to own cows and other livestock, so they sold her property (1 cow and part of her land for Ksh. 60,000) to cater for the costs of her funeral*	
	2	**Borrowing and/or taking loans from family members, neighbors, friends and local table banking groups**[160] in order to supplement funeral costs.	*We borrowed 500 Ghana cedi for the funeral; I took 200 of that from the VSLA [a table banking group]* *Because of my mother's loss, I borrowed KES 20,000 three years ago and have not refunded the money until today*	

[160] Table banking groups are group funding exercises where members of local interest or friend groups meet and contribute a specified amount of money every week/month/year. Different members of the group can then either receive the lump sum every week/month/year or borrow under specified interest rates.

	3	**Depleting all the household savings** or working extra shifts in order to get enough money to contribute towards the funeral expenses.	*I personally contributed KES 3000, money I had saved and that I was planning to use on buying a cow* *My family lost a lot of money to support the funeral—my husband spent around Ksh.5000 on the burial (this is all the money he had in his savings). My father-in-law had just received Ksh.8000 at the time from his elderly cash transfers, and he used all of it in the funeral expenses*	
	4	**Receiving money, assets** (such as livestock and farm produce) **and services from community members** in order to cover funeral costs.	*The money that was spent on this funeral (KES 100,000) was from contributions from the community members* *The community contributed money to the total funeral expenses of around KES 210,000*	
	5	**Receiving money from associations in which the deceased was a member** in order to support the family of the deceased.	*[The deceased] used to be a motorbike rider so the association of motorbike riders paid for all his funeral expenses; they even gave his wife Ksh. 45,000 to support them for a while*	
	6	**Spending time away from income-generating activities** during funeral activities and while grieving. As a result, respondents note losing farm produce and regular income.[161]	*The grandparents of my child; his mother and I were affected; we grieved for 3 months before we finally let go. We couldn't go to farm, and as a result, lost some of our farm quotas and bambaram beans* *I did not work for 1 week before the burial and 3 days after the burial because I was preparing the ground for burial. After the burial, I still stayed home as per the cultural requirements. I therefore lost income during those 10 days*	
	7	**Losing jobs from inflexible employers** upon returning from funerals of loved ones due to the high time costs of funerals.	*In one of the families [of the funerals I attended], one of the members who lost her mother spent three weeks mourning and by the time she went back to work, she found that her employer had sacked her. She has not been able to find a job again* *My sister was rendered jobless since she overstayed at home during the funeral. Her employer terminated her employment for this reason. She stayed for 1 week at home*	

[161] For many respondents, funeral exercises take anywhere between one week and three months. In addition to the time spent of grieving and fulfilling responsibilities associated with planning the funeral, respondents mention that socio-cultural stipulations require people stop working until the deceased is buried.

Long-term economic effects	8	**Losing current and future socio-economic resources** upon the death of children and adults respectively.	*I lost my regular remittances because of my son's death* *The man who died was the most successful in the family and other families used to go to him to solicit for funds and fees, so they've also been left hungry; he used to make sure that none of the young boys were wanderers – he would bring them in and teach them carpentry skills; now that he's passed, that has all been transferred to my husband. He is trying*	
	9	**Caring for dependents of the deceased** by taking economic, emotional and overall social responsibility for those left behind.	*I was really sad for a long time, especially since I had lost my daughter and had to take care of the two children who had been left behind* *She died during childbirth in a government hospital… she died after being in the hospital for three days. The child survived and now lives with us*	
	10	**Experiencing adverse effects on education** as a result of losing primary providers.[162] These include dropping out of school and, as a result, being involved in early marriages.	*[He] died and a month later the wife died so there was quite a lot of sadness. The other relatives took up the children who were young. There was a lot of sadness, and the relatives who took up the children were not able to fully support them so 2 children stopped going to school and got married so as to ran away from the financial troubles* *I dropped out of school because my father passed away—my mother had also passed away years prior*	
	11	**Inheriting land, houses and property of the deceased**[163] ensures that the family members who are left behind have socio-economic support. *	*We are doing well now because we inherited the remaining farm, cows etc. from the deceased* *She did not leave too much behind apart from her land, which was now passed on to the son*	

[162] This finding aligns with Evans and Miguel 2007 in their analysis of a 5-year panel data set on school attendance after a parent's death. They find that school participation falls 5.5 percentage points after parent death.
[163] This applies to cases where the deceased was in a different household. e.g. an old female leaving her property for her son and his household.

APPENDIX 8: AGGREGATED RESULTS FROM THE PRIMARY METHODS

APPROACH TO AGGREGATING RESULTS

As we collected data across three primary methods, and regarding more age groups than just under 5 and 5 or older, several steps were required to make our results directly comparable to GiveWell's moral weights:

1. For each of our three primary methods, we form weighting averages according to estimated age distribution among the relevant population to form estimates for individuals under 5 and individuals 5 or older.
2. For the community perspective, we take the ratio of under 5 to 5 or older from our relative value of different ages choice experiment and use this to convert the under 5 value from our monetary value of life choice experiment, to an equivalent value for individuals 5 or older.[164]
3. Finally, we take a weighted average of the individual and community perspective to give a single under 5 and 5 or older value, which can then be converted to moral weights.

However, we recognize there is substantial subjectivity in this process. We made a number of decisions based on our best technical judgements, but acknowledge the final aggregated results are sensitive to these decisions and that GiveWell staff members may have differing views about the best approach. Below we outline the main decisions we made, and illustrate the sensitivity of the final aggregated results to each decision.

Decision 1: Individual and community perspective are of equal relevance

As noted above, we make a default assumption that both perspectives are of equal relevance to moral weights. GiveWell staff members may want to place more weight on one perspective (see Table 5 in Results Overview for comparison; this assumption can be adjusted in the accompanying Aggregating Results excel sheet). To demonstrate how this changes the outputs, in Table 40 how the aggregated results vary if more weight is put on each perspective (we've arbitrarily used 80%).

Table 40. Sensitivity of aggregated results to relevance weighting between the individual and community perspective.

	Default equal weighting		80% weight to Community Perspective		80% weight to Individual Perspective	
	Value (USD)	**Moral weights**	**Value (USD)**	**Moral weights**	**Value (USD)**	**Moral weights**
Individuals under 5	$65,906	230	$81,058	283	$50,837	178
Individuals 5 or older	$40,721	142	$44,644	156	$36,510	128

Decision 2: Not to incorporate confidence weighting in the default result aggregation.

In the pre-analysis plan, we specified that in addition to weighting our methods by relevance, staff members likely want to incorporate weighting by confidence. We suggested forming a weighted

[164] For simplicity, we assume independence across different estimates, and use simulations to obtain the distribution of the product.

average, which linearly penalizes methods for the number of respondents dropped from the final estimation. For VSL 63% of respondents are included in the final sample; 37% are dropped for failing to pass the internal scope test, or failing to understand at least one small probability test question after 2 or more explanations. For the community perspective choice experiment, 89% of respondents are included in the final sample; 11% of respondents are dropped for failing the dominance test or exhibiting clearly inconsistent preferences. Following the protocol laid out in the pre-analysis plan, this would lead you to place slightly more weight on the community perspective.

In our final aggregated result, we have not included this confidence weighting. Despite trying to make the tests as comparable as possible, it remains easier to apply understanding tests to VSL than to our choice experiments. As we test understanding in more ways, there are more opportunities for respondents to fail test questions; we are concerned that this would lead to undue bias away from VSL results. Furthermore, the weak external scope test, typically used in the stated preference VSL literature as a key test of validity,[165] involves an arbitrary cutoff at the p-value of 0.05, which seems undesirable.

The table below demonstrates the impact on the aggregated results if the confidence weighting suggested in the pre-analysis plan is used.

Table 41. Sensitivity of aggregated results to including a measure of confidence in the weighting across approaches.

	Default equal weighting		Confidence weighting used	
	Value (USD)	**Moral weights**	**Value (USD)**	**Moral weights**
Individuals under 5	$65,906	230	$69,569	243
Individuals 5 or older	$40,721	142	$41,324	144

Decision 3: Which central estimate to use (whether to exclude over 40 valuation from community perspective)

As we've shown in the report, across our methods there is some sensitivity in our results to which analytical model to used. Throughout, we have stuck with the analytical model specified in the pre-analysis plan to avoid bias in the selection of which model to use.

The exception to this, is with the results of our community perspective, relative value of different age groups, choice experiment. As described in more detail in Section 2, our estimation model outputs a highly negative valuation for individuals over 40. When included in our calculation of the under 5 to 5 or older ratio this has a nontrivial effect on the results. As such, for our aggregated results we chose to use the ratio that excludes those over 40. GiveWell staff members may decide instead to use the aggregated result including this age group. As shown below, this results in substantially lower weight being placed on individuals aged 5 or older.

[165] For instance, it is a criterion used by the US Environmental Protection Agency to decide whether to incorporate the results of a stated preference VSL study into its estimate (US EPA 2016). https://yosemite.epa.gov/sab/sabproduct.nsf/0/0CA9E925C9A702F285257F380050C842/$File/VSL%20white%20paper_final_020516.pdf

Table 42. Sensitivity of aggregated results to the inclusion of very low valuation of individuals over 40 in the community perspective.

	Default - excluding over 40		Including over 40	
	Value (USD)	**Moral weights**	**Value (USD)**	**Moral weights**
Individuals under 5	$65,906	230	$65,947	231
Individuals 5 or older	$40,721	142	$29,282	102

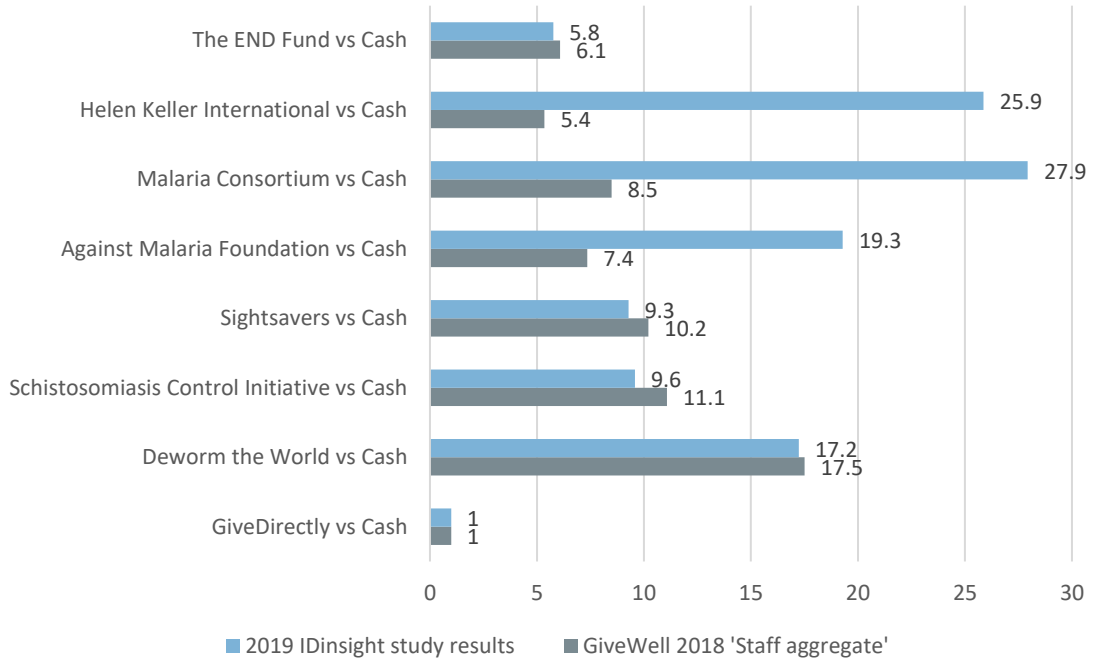
Decision 4: $286 is the appropriate value for conversion between USD and moral weights

As our raw results are presented in USD, a conversion must be made into the moral weight unit, to allow comparison of our values with GiveWell's current values. To make this conversion we take the 'value of doubling consumption for one person for one year' to be equivalent to the average annual consumption per capita assumed for the typical beneficiary population throughout the GiveWell CEA model (286 in nominal USD). If we hold the value of doubling consumption as 1, this effectively gives us the value in USD of 1 unit of moral weights. So, we divide our values in USD by 286 to estimate moral weights based on our data from beneficiaries.

IMPACT ON CEA MODEL

To estimate the effect of incorporating beneficiary preferences on the GiveWell CEA results, we inputted our aggregated moral weights into the latest version of the model available online. As shown below (Figure 17) the most notable impact is the substantial increase in the cost-effectiveness ratios for Against Malaria Foundation (2.6 times increase), Helen Keller International (4.8 times increase), and Malaria Consortium (3.2 times increase).

Figure 17. Comparison of resulting cost-effectiveness ratios for top charities - GiveWell 2018 'Staff Aggregate' vs IDinsight 2019 study results.



APPENDIX 9: FINDINGS FROM OUR SECONDARY METHODS

Following piloting, we had the highest level of confidence in the selected three primary methods to reliably capture preferences. However, given the difficulty of measuring such complex preferences, we still had some uncertainty over our approach. This was in particular true of the choice experiment to measure the monetary value of life, as we have not seen a similar approach previously in the literature, so cannot compare our results to external data. As such, we included two secondary methods for a random 10% subsample of respondents as a robustness check on our main results. Both secondary methods are direct variations on that community perspective choice experiment, as summarized in Table 43.

Table 43. Summary of secondary methods

Method	Comparison to primary choice experiment	Additional information provided
Relative value of life and education	In place of cash transfers, the second item offered by each program is education to the end of secondary school.	1) Alternate approach to estimate an implied value of life – robustness check on main approach, 2) Some data on how respondents value a different outcome (education, instead of cash).
Monetary value of life (one-sided)	Rather than two programs that both offer two goods at different levels, this presented a direct trade-off between a program that saves lives, and another that gives cash-transfers.	1) Alternate approach to estimate relative values of different age groups – robustness check on main approach, 2) Tests the theory that respondents are less likely to engage with such directly presented trade-offs (as seen throughout piloting).

We have low confidence in the monetary values obtained from our secondary methods and do not suggest these are used directly by GiveWell. This is due to the non-monetary values respondents attach to education, the additional biases created when using a direct (one-sided) framing, and the smaller sample size. However, **we found that the derived values are within the same range as those from our primary values, and once again we find that highest value is placed on individuals under 5. This gives further confidence to the high-level findings of our main methods.**

RELATIVE VALUE OF LIFE AND EDUCATION

TAKEAWAYS

- Across the entire sample, **we see a consistent pattern of having more life non-switchers in Ghana compared to Kenya** (as in the monetary value of life results in Section 1).
    - This reassures us of the consistency of respondents' preferences across the primary and secondary methods.
- Methodologically, **it is difficult to convert the relative value of life and education to a monetary value** due to the huge heterogeneity in the estimated returns on education as well as the non-monetary values placed on education.
- However, **using this approach we estimate that the monetary value of life is between $38,261 and $114, 782,**
    - This is within a similar median range as the results in the monetary value of life choice experiment described in Section 1.

This section details the results of beneficiaries' valuation of life relative to education from data we collected with a random sample of 10% of respondents (n=160). Like cash transfers, education can result in better employment opportunity and associated increases in long term income. In theory this gives us an alternate conversion unit between lives saved and dollars, assuming we can estimate the returns on education in dollars.

We initially considered this approach due to concerns that respondents may be undervaluing the return from cash transfers during our piloting phase. However, we see this as a supplementary method because we found during piloting that people often place non-monetary values on education that are difficult to measure and quantify. We are able to estimate the expected value of an increase in income that an individual expects with secondary education, but the value of education may be inflated by other gains (such as status within the community) that we cannot measure in dollars.

Despite our lower confidence in the results from this method, it provides a useful consistency check on our other choice experiment results.

Overview

In this method we:

1. Started with a dominance test in which respondents needed to choose between two programs where one is clearly inferior.
2. Ask respondents for their perceived economic returns to primary and secondary education.
3. Conducted a choice experiment in which respondents trade-offs between two hypothetical interventions. Each intervention saves the lives of a number of children under 5 and provides secondary education to a number of children who would otherwise only complete primary education.

Each respondent was given two questions total, where the second question depended on their response to the first. The first question was the same for each respondent, as follows:

"Program A saves the lives of 5 children aged 0-5 years AND supports 10 children throughout the end of high school. Program B saves the lives of 6 children aged 0-5 years AND supports 5 children throughout the end of high school. Which one would you choose?"

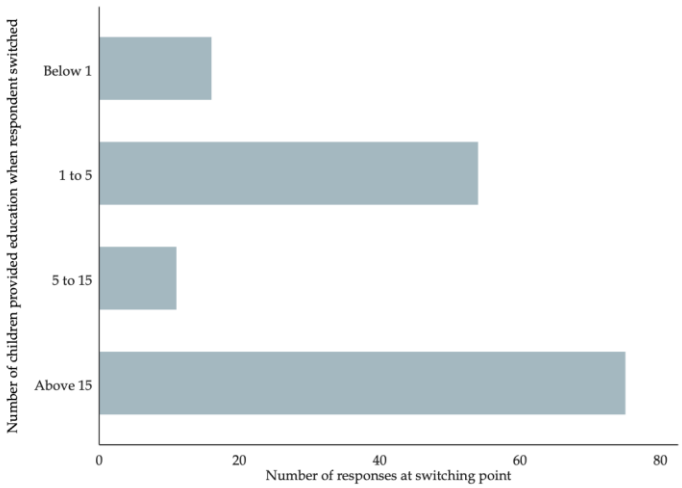
If they chose Program A, we changed the choice to Program A saving 5 children and providing education for *6* children while Program B remained unchanged. Otherwise, we changed the choice to Program A saving 5 children and providing education for *20* children while Program B remained unchanged.

Results

We found good evidence that respondents understood the trade-offs presented. Across the entire sample, nearly all (97.5%) passed the dominance test, and 100% demonstrated consistent preferences across their choice set.[166]

Overall, **48% of respondents chose saving a life over providing education for an additional 15 children**. The median switching point for the full sample is between 5 and 15 children who are provided education till the end of secondary school. Similar to our primary methods, **Ghana has a higher fraction of respondents always choosing life**. The median switching point for the Ghana sample is above 15 children, and the median for the Kenya sample is between 1 and 5 children. Below we show the distribution of choices, illustrated by the "switching intervals" at which people switch from choosing the program with more lives saved to the one offering education to more children.[167]

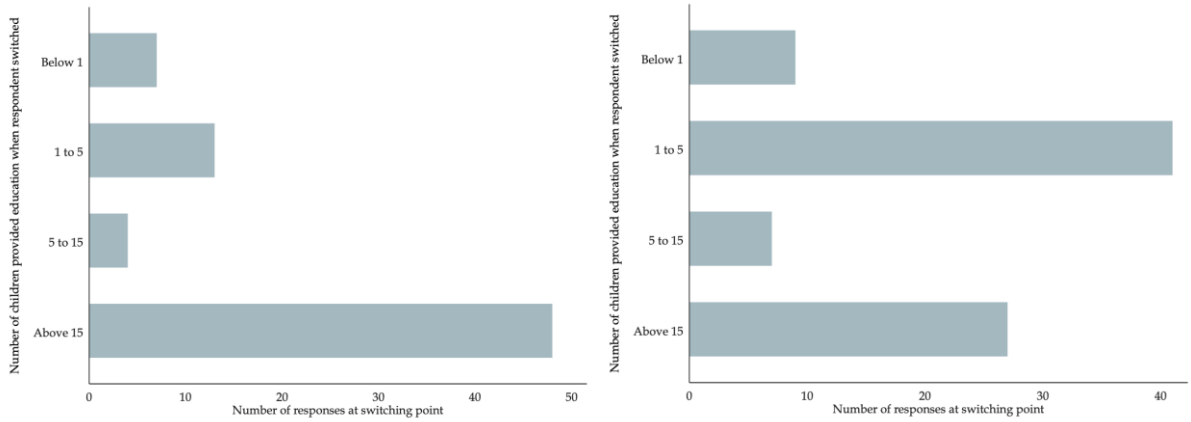
Figure 18. Distribution of education vs. life choices across the entire sample (N = 156)



[166] It is important to note that consistency in this method was measured across 2 choices. This makes it easier to pass the consistency check compared to our other choice experiments where more choices were presented.

[167] These graphs show the intervals in which a respondent switches from choosing to save one child (when the number of children receiving education is below the lower bound of the interval) to choosing to give education to a number of children (when the number receiving education is within the interval). For instance, "Below 1" means the respondent chooses to give education even when it's one child saved and one child educated. "1 to 5" means the respondent chooses to give education to one to five children instead of saving one child. "5 to 15" means the respondent chooses to give education to five to fifteen children instead of saving one child. "Above 15" means the respondent chooses to give education to more than 15 children instead of saving one child.

Figure 19. Distribution of education vs. life choices across respondents from Ghana (left; N = 72) and Kenya (right; N = 84)



In order to come up with a central estimate for the monetary value of life from the education point estimates, we captured perceived economic returns to education. We asked the following questions:

1. Typical careers one has after completing primary or secondary education,
2. Estimates of the average durations spent in these careers, [168] and
3. Annual incomes from these careers[169]

We estimated the average expected return on education as $7,652 across our full sample.

To estimate an implied monetary value of life, we take our median switching point (between 5 and 15 children educated) and multiply by the estimated returns. **We estimate that the implied monetary value for averting a death is between $38,261 and $114, 782 using this approach**. We noted large heterogeneity in the estimates of returns on education. Values for Kenya were considerably higher, and driven by a small number of outliers. To reduce the impact of these outliers we have used the average estimate of return on education throughout, as we feel this gives more reliable final estimates. Using country level, or even individual level, estimates gives a much larger range of values.[170]

[168] For simplicity, we asked for the age at which one can start working, and subtracted it from 67 (Kenya's life expectancy) to arrive at the duration. On average, respondents believe someone with completed primary school education could start working at age 20, and secondary school 23.

[169] On average, respondents think a primary school graduate will make 362.9 USD a year, and a secondary school graduate will make 1271.5 USD a year. However, Kenya averages are much higher than Ghana's, and Kenya's distribution has very high outlier values (the averages are 111.5 USD and 195.5 USD in Ghana and 610.5 USD and 2347.5 USD in Kenya for primary and secondary school respectively).

[170] If we instead use an individual estimation approach, where we take an individual's estimated switching point (taking the midpoint of their switching intervals) and an individual's estimated return on education, the mean value of averting death is $444,713 for the full sample ($39,328 for Ghana, and $837,391 for Kenya). The median using the same approach is $16,505 for the full sample ($8,915 for Ghana, $40,369 for Kenya) demonstrating how sensitive these estimates are to high outliers.

Table 44. Aggregate estimates for the monetary value of life using education vs life choice experiment.

	Median switching point	Median implied value of averting death (USD)
Full sample	5 to 15	38,261 to 114,782
Ghana	>15	> 114,782
Kenya	1 to 5	7,652 to 38,261

Comparison to cash vs. life choice experiment

Comparing the results from our aggregate approach to the results from the cash vs. life choice experiment, we see that the median estimates are roughly consistent across both methods.

Table 45. Comparing results from the aggregated approach to results from the cash vs. life choice experiment.

	Approximate Median (USD): education vs life	Approximate Median (USD): cash vs. life
Full sample	38,261 to 114,782	80,000
Ghana	> 114,782	> 95,000
Kenya	7,652 to 38,261	30,000

Overall, this choice experiment exhibits consistent overarching results to the cash vs. life choice experiment. This includes the significant presence of "non-switchers" and the consistent country-level pattern of having more pro-life non-switchers in Ghana compared to Kenya. This reassures us of the consistency of respondents' preferences across different elicitation methods. Further, our results corroborate the finding from piloting that there is large heterogeneity among respondents in their views on different types of programs or outcomes (demonstrated by the wide range of estimates for the returns on education). This warrants future research to better understand the diversity of preferences across different outcomes.

MONETARY VALUE OF LIFE: ONE-SIDED CHOICE EXPERIMENT

TAKEAWAYS

- The relative valuation of different age groups was consistent with the findings in Section 2. **Individuals under 5 are consistently valued higher than individuals 5 and older** (with an approximate ratio of 1.26).
- **In the one-sided approach, 48% of respondents who always to save the life of a single child under 5,** compared to 38% in the two-sided approach (as discussed in Section 1).
- The proportion of respondents who always choose life in the one-sided approach is slightly higher, suggesting preferences are sensitive to the directness of the framing used.

This section details the results of beneficiaries' valuation of life and cash using a one-sided framing approach, captured from a random sample of 10% of respondents (n=180). Instead of having respondents choose between two programs, both of which save some lives and give some cash transfers, we directly compare a program that saves a life with a program giving a number of cash transfers. When we used this method in the pilot, we found that a high proportion people would always choose to save the life even at very high cash transfer values. As such, we adopted a "two-sided" version for the main monetary value of life choice experiment which led to decreased levels of non-switchers during piloting.

We believe that the directness of the framing of the question led to extreme responses, an effect that has been observed elsewhere in the literature. People are often influenced by "reputation concern", the perception of their choices by others whenever they make moral judgments (Lee et al. 2018, Bénabou et al. 2018). In particular, Lee et al. find that reputation concern can increase deontological views in moral judgments (consistent with our piloting results). Since the one-sided approach has a more direct and repugnant trade-off compared to the two-sided approach, we hypothesize that the "reputation concern" is higher. This could lead to more deontological views on the inherent value of life, explaining the higher proportion of respondents who always choose life in the one-sided approach.

Overview

Here we discuss the results from the "one-sided" choice experiment to provide further evidence on the relative values of age groups, and to test whether framing had an effect on results. The question framing was as follows:

> *"Suppose a donor is choosing between two options: buying a medicine that costs X USD which can be used to save the life of a Kenyan aged Y who would otherwise die from a disease, and giving cash transfers worth X to extremely poor Kenyan for them to improve their lives. Which one do you think the donor should choose?"*

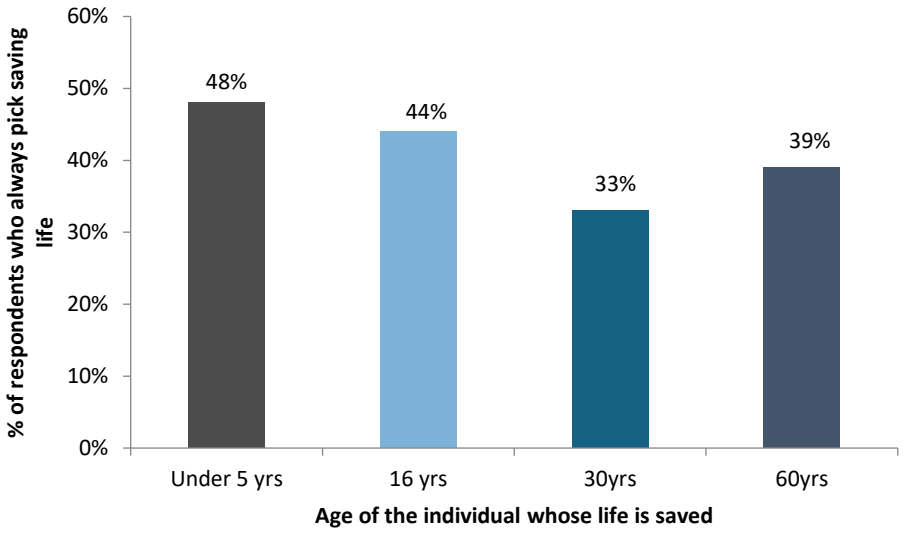
We varied the value of *X* incrementally from $1,000 to $10,000,000 and randomized *Y* between individuals aged under 5, 16, 30 and 60. We asked about individuals under 5 years to all 180 respondents, and other age groups (16, 30 and 60) to approximately one third of the 180 respondents each (to decrease the effect of survey fatigue).

Results

Across the entire sample, **the percentage of respondents choosing to always save a life instead of giving cash transfers is highest for individuals under 5 years**—48% of respondents always pick saving the life of an individual under 5 years instead of giving cash transfers. This percentage declines as we go towards older age groups, except for 60-year-olds, as shown below. Given the small sample size (n = 74), we have more confidence in the main relative value of age groups choice experiment, where we obtained the following ranking: under 5 years > 5-18 years > 19-40 years > over 40 years. However, it does provide further evidence that the negative valuations for individuals over 40 produced through the main approach were a result of a problem with the estimation model rather than a true preference.

The graph below shows the distribution of respondents who always picked the life-saving program at different ages. Based on these percentages **we find an approximate ratio of 1.26 between the relative valuation of under 5 and 5 and older lives, which is consistent with the findings in Section 2**.

Figure 20. Distribution of respondents who always picked the life-saving program at different ages



As expected, we find the proportion of respondents always choosing to save an individual under 5 is higher with this more direct framing (48% non-switchers) than with our two-sided approach (38% non-switchers). Again, we found substantially higher levels of non-switchers in Ghana than in Kenya.

Below we present graphs of switching patterns across the different age groups.

Figure 21. Switching patterns when picking between saving an individual under 5 versus giving cash transfers (left; N = 180) and saving an individual 16 years old versus giving cash transfers (right; N = 55)

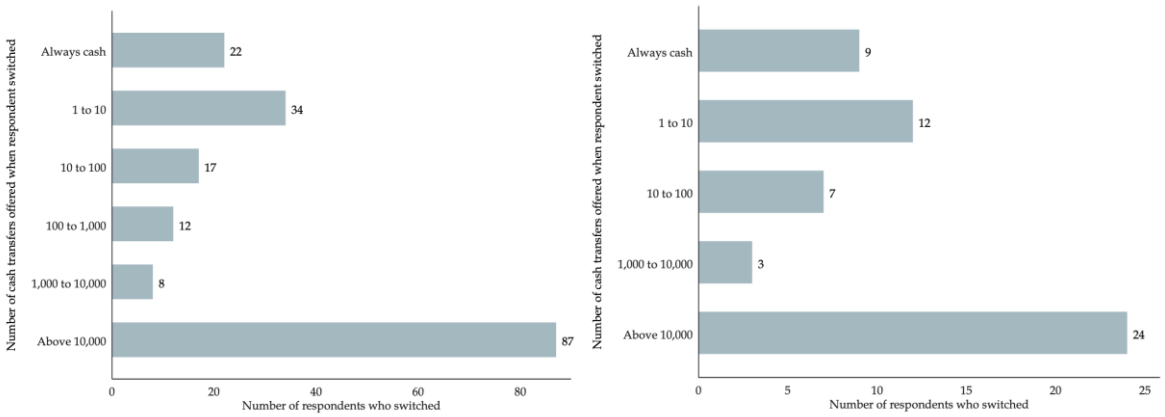
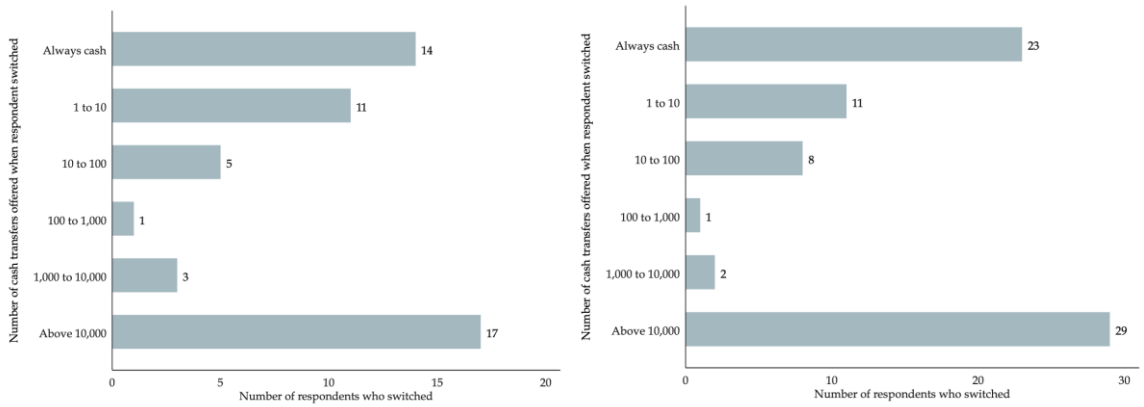


Figure 24. Switching patterns when picking between saving an individual 30 years old versus giving cash transfers (left; N = 51) and saving an individual 60 years old versus giving cash transfers (right; N = 74)



APPENDIX 10: STUDY LIMITATIONS

Our study, especially the survey on value of statistical life, employs state of the art techniques in the literature adapted to local contexts after extensive piloting to maximize respondent understanding. However, we would like to point out a number of inherent limitations of our approach:

- Representativeness: For the VSL estimation, we include only respondents with sufficient understanding, which is about 62% of all respondents. They differ in a number of demographic variables from those who did not exhibit sufficient understanding (see Table 19). We account for these differences in the inverse probability weighting adjusted estimates (which changes results by 3%: from 35732.62 for 5/1000, model A, to 36798.58), but we are not able to account for any differences driven by unobservables which could still lead to bias in results.
- Liquidity constraint: As discussed, we find evidence that the ten-year small installment repayment framing helped relax liquidity constraint, but cannot rule out that within each period respondents' payment amounts are constrained by liquidity.
- Scope sensitivity: The WTP values are far from proportional to the risk reduction levels (failing the strong external scope test), even though the values pass the weak external scope test. It is not clear to what extent we should expect them to be proportional since this is a theoretical result that applies to "small" probabilities, and studies from high-income countries also often fail the strong external scope test. However, this does pose a challenge for validity of the result and extrapolating the VSL from one value of "small" probability to another (e.g. from 5/1000 to 10/1000).
- Hypothetical bias: Since the survey is based on stated preference (we have piloted a revealed preference-based approach which turned out to be infeasible) it inevitably suffers from potential hypothetical bias. US EPA's review of VSL studies in the US finds the mean across hedonic wage (revealed preference) VSL studies is 11.9 million USD, and that across stated preference studies is 8.6 million USD, though the revealed preference studies have the issue that people may not know the risks precisely and hence are less than perfect in capturing true preferences. Given this, it is possible that our estimates are somewhat lower than true values, but the direction and magnitude of the bias are not entirely clear. The same criticism applies to the choice experiments, which may suffer from even more hypothetical bias as allocating resources among age groups or between interventions is less familiar to people than trading off between income and mortality for oneself and one's family.
- Lack of robustness of choice experiments estimates: Data from choice experiments on community perspective "moral views" on the relative values of age groups and the trade-off between cash transfers were analyzed using the standard logit model assuming homogeneous preferences. However, due to large heterogeneity in preferences observed in our context, as well as lack of switching for certain pairs of choices, some of the estimates (e.g. that for those aged above 40, and most estimates on the "cash-vs-life" choice experiment) exhibit negative values which are unintuitive and unlikely to reflect true preferences.

We believe it is important to recognize the value of the study — namely being one of the first studies to systematically estimate the preferences among aid beneficiaries over different outcomes — while accounting for its limitations when applying the results.