USAID | UGANDA
FROM THE AMERICAN PEOPLE

SAVING MOTHERS, GIVING LIFE PHASE I IN UGANDA: COST EFFECTIVENESS ANALYSIS

February 6, 2016

This publication was produced for review by the United States Agency for International Development. It was submitted by the USAID/Uganda Monitoring, Evaluation and Learning Program, AID-617-C-13-00007 It was prepared by SoCha,llc under subcontract to QEDGroup,llc.

SAVING MOTHERS, GIVING LIFE PHASE I IN UGANDA: COST EFFECTIVENESS ANALYSIS

Submitted to:

USAID/Uganda

Prepared by:

SoCha,llc, under subcontract to QEDGroup,llc

Disclaimer:

The authors' views expressed in this document do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

Acknowledgements

The debts we owe in regards to this report are many. First, we would like to thank USAID/Uganda for advancing this innovative approach to cost effectiveness analysis. In particular, we owe special gratitude to Lesley Reed, Mark Messick, Anne Murphy, and Margaret Sancho for the strong support they provided to this exercise, as well as to Janex Kabarangira and Patricia Mwebaze for the assistance they gave us building strong relationships with the implementing partners. We would also like to thank Mark Tumwine and Walter Obiero from CDC/Uganda, whose assistance with obtaining crucial datasets paved the way forward for this analysis. Staff from USAID/DC and CDC Headquarters – including Claudia Morrisey, Florina Serbanescu, Mona Steffen, and Paul Stupp – provided useful feedback during our various findings presentations that helped sharpen our focus.

On the implementing partner side, we would like to thank Baylor staff members Dr. Adeodata Kekitiinwa, Dr. Dan Murokora, Dr. Alice Assimwe and Wycliffe Ssewankambo in Kampala, and Dr. Leonard Ssenyonjo, and Vincent Kamara for their incredibly kind assistance in the field. On the IDI side, we would also like to thank Dr. Alex Coutinho, Brenda Picho and Agnes Nalutaaya in Kampala, and Dr. Joanita Kigozi, Alex Muganzi and Charles Mukalazi in Kibaale. We also offer gratitude to Dr. Margaret Kruk and Daniel Vail from Columbia University for sharing select datasets. We benefited greatly from their previous evaluative efforts of the program.

QED and the Learning Contract Team, including Paul King, May Mwaka, Lane Pollack and Rand Robinson were supportive from the beginning. We also owe a special gratitude to Johan Kleinhans, who drove the effort forward and laid the essential groundwork for this to proceed under QED. Without his support, this exercise would not have been completed. Finally, the support and guidance we received from Phillip Greene, USAID/Uganda Economist, was absolutely crucial. Any mistakes in this report should in no way detract from the valuable assistance they provided, and remain solely our responsibility.

-SoCha

Acronyms

BEmONC	Basic Emergency Obstetric and Neonatal Care
CDC	U.S. Centers for Disease Control and Prevention
CEA	Cost Effectiveness Analysis
CEmONC	Comprehensive Emergency Obstetric and Neonatal Care
CLA	Collaboration, Learning and Adaptation
DHMIS	District Health Management Information System
EMK	Essential Medicines Kit
EmONC	Emergency Obstetric and Neonatal Care
GoU	Government of Uganda
HFA	Health Facility Assessment
IDI	Infectious Diseases Institute (Makerere University, Uganda)
JMS	Joint Medical Stores
M&E	Monitoring and Evaluation
MMR	Maternal Mortality Ratio
NMS	National Medical Stores
OGAC	US Office of the Global AIDS Coordinator
PI	Principle Investigator
POMS	Pregnancy Outcome Monitoring Survey
QCA	Qualitative Comparative Analysis
QED	QEDGroup,LLC
SDS	Strengthening Decentralization for Sustainability
SMGL	Saving Mothers, Giving Life
SoCha	SoCha,LLC
UGS	Ugandan Shillings
USAID	United States Agency for International Development
USD	United States Dollar
USG	United States Government
VHT	Village Health Team

Table of Contents

Acknowledgements ii
Acronyms iii
Table of Contents iv
Executive Summary vi
Findings on how to reduce MMR vii
Cost Model Findings viii
Cost Effectiveness Results ix
Recommendations x
Program Design Recommendations x
M&E Recommendations based on conducting the CEA xi
Background 1
Rationale for this study 2
Scope of Work 2
Defining CEA 3
Level of Effort, Unit of Analysis and Data Collection 4
Objective 1: Identify which combinations of SMGL Phase 1 activities are necessary and/or sufficient to improve maternal survival in Uganda 7
Findings 8
Explaining the Findings 10
Level III and Infrastructure Pathways 10
High Demand and Kibaale/Ambulance Pathways 14
Private Facility Pathway 16
Conclusion 18
Objective 2: Build a model that can project the approximate costs of implementing SMGL Phase 1 activities elsewhere in Uganda 19
Findings 21
Conclusion 22
Objective 3: Combine the above listed analytical components into a third that compares the relative costs of each combination of activities to determine which are the most cost effective 23
Findings 23
Wider Applicability of the CEA Findings and Their Reliability 24
Conclusions 24
Recommendations 26
Program Design Recommendations 26
Recommendations based on Lessons Learned from Using Program and Outcome Data for Model Building 27

Annex 1: Scope of Work i
Annex 2: QCA Background and How We Modeled SMGL iii
How We Modeled SMGL Phase 1 v
Step 1: Model the MMR Outcome and the SMGL Intervention into a Truth Table vi
Step 2: Conduct Exploratory Analysis to Eliminate Irrelevant Factors, Expand Coverage, and Refine the Model(s) xv
Step 3: Conduct Confirmatory Analysis to Identify and Consolidate Pathways xvi
Annex 3: How We Modeled SMGL's Cost xxv
Costing Methodology and Model Design xxv
Structure of Cost Categories xxvii
Data Sources xxx
Site Visits, Interviews and Collection of Additional Information xxxi
Limitations xxxii
Annex 4: Cost Model Instruction Manual xxxiii
Basic Concepts for the User xxxiii
Protected Sheets, hidden rows and secret numbers xxxiii
Colour Formatting - Yellow versus Blue xxxiii
Drop Down Menus xxxiii
Example One: Standard Sheet Layout xxxv
Example Two: Personnel Assumptions and Job Titles xxxvi
Explanations of the Sheets xxxvi

Executive Summary

Saving Mothers, Giving Life (SMGL) is a 5-year initiative designed to rapidly reduce deaths related to pregnancy and childbirth through a coordinated approach that strengthens maternal health services in high-mortality settings. SMGL focuses on reducing the "three delays" that prevent women from accessing maternal health services: delays in seeking services, delays in reaching services and delays in receiving quality care. These interventions are focused primarily on the critical period of labor, delivery and the first 48 hours postpartum, when most maternal deaths and approximately half of newborn deaths occur.

Focusing on four districts in western Uganda (as well as in Zambia), SMGL's Phase 1 took place from July 2012 through June 2013. Altogether, 107 Ugandan health facilities that provided delivery care received support from Baylor College of Medicine Children's Foundation Uganda (hereafter Baylor) in the Kabarole, Kyenjojo and Kamwenge districts, and from Makerere University's Infectious Disease Institute (IDI) in the Kibaale district. In close collaboration with the Government of Uganda (GoU), SMGL Phase 1 supported 24 different evidence-based interventions in the community and in health facilities, including:

- Equipping facilities to provide high quality emergency obstetric and newborn care (EmONC) in a geographically strategic way, enabling women with complications to receive care within two hours.
- Improving supply systems so facilities have the equipment, supplies, commodities and drugs needed to deliver high quality EmONC services.
- Training and mentoring skilled birth attendants to provide quality, respectful delivery services, and stabilize, treat and refer emergency cases if necessary.
- Mobilizing the community, through Village Health Teams (VHTs) and other mechanisms, to generate demand for facility deliveries, antenatal and postpartum care while emphasizing birth planning, and encouraging HIV testing and treatment and uptake of family planning services.
- Strengthening linkages between communities and facilities through a 24/7 integrated communications and transportation system, including the use of subsidized vouchers for motorcycle (Boda) transport[1] and free deliveries at private facilities, to help pregnant women access childbirth facilities in a timely manner.
- Improving data collection systems to record pregnancy outcomes, including complications and maternal deaths, and strengthen host country health management information systems.

SMGL activities included routine monitoring during the first year of the project, including baseline and endline assessments at the beginning and end of Phase 1.

SMGL's Phase 1 results are impressive. Drawing upon an extensive collection of outcome data, the Centers for Disease Control (CDC) found that the maternal mortality ratio (MMR) per 100,000 births fell by at least 30% over the course of the 12 months. This reduction was driven by women's increased access to and receipt of emergency obstetrics and newborn care. These and other findings reveal that Phase 1 delivered remarkable results, contributing to a rapid decline in the number of women who die during pregnancy and childbirth in SMGL program areas in Uganda (as well as in Zambia).

Plans are underway to move SMGL into Phase 2 by expanding to other districts in Uganda and Zambia (as well as Nigeria) with a long-term goal of national scale-up to save more women's and newborns' lives. However, pressing questions remain unanswered: Does Phase 2 need to deliver all 24 activities to reduce MMR by similar levels elsewhere, or can SMGL be as effective with less (i.e. a so-called "essential package)? If so, which combinations of activities are the most cost effective?

QEDGroup,llc (hereafter QED) subcontracted SoCha,llc (hereafter SoCha) through the Uganda Monitoring, Evaluation and Learning Contract (aka The Learning Contract) mechanism to perform a Cost Effectiveness

[1] The Boda Voucher program was funded separately by the foundation Every Mother Counts, and was implement by Baylor through VHTs.

Analysis (CEA) of SMGL Phase 1. The overall aim of the CEA was *to determine which aspects of the Phase 1 SMGL intervention are the most cost effective in preparation for a Phase 2 rollout.* The CEA's primary objectives were to:

- Identify which combinations of SMGL Phase 1 activities are necessary/sufficient to improve maternal survival in Uganda;
- Build a model that can project the approximate costs of implementing SMGL Phase 1 activities elsewhere in Uganda; and
- Combine the above-listed analytical components into a third that compares the relative costs of each combination of activities to determine which are the most cost effective.

To meet these objectives, SoCha combined a "qualitative comparative analysis" (QCA) approach with cost modeling to conduct the CEA.

Findings on how to reduce MMR

Informed by QCA, the SoCha team identified five program pathways (combinations of SMGL activities) that are sufficient to reduce MMR. The first four pathways are specific to government health facilities and identified according to the context in which they best fit. The final pathway is associated with private health facilities (for profit and not-for profit):

- ***Pathway 1: The Level III Pathway, combining Boda Vouchers & Mentoring & Medicines (and assumes VHTs are present)***
  Combining a Boda transport voucher option (which assumes VHTs are present) with mentoring of staff, and including the provision of essential medicine is sufficient to reduce MMR by at least 30%. This pathway runs across Level III health facilities with moderate increases to demand for facility-based births (less than 100%).

- ***Pathway 2: The Infrastructure Pathway, combining Infrastructure & Boda Vouchers & Mentoring & Medicines (and assumes VHTs are present)***
  Pathway 2 is a subset of Pathway 1. In facilities that require new or rehabilitating operating theaters, offering the same combination of the provision of critical medicines, Boda Vouchers (which assumes VHTs are present) and ongoing mentoring of critical staff is sufficient to reduce MMR by at least 30%. This pathway is also effective at Level IV government facilities that require new or upgraded or operating theaters.

- ***Pathway 3: The High Demand Pathway, combining Ambulances & Medicines & Equipment & Staff***
  Combining the ongoing provision of essential medicines with ambulance transport services, the provision of essential medical equipment and additional staff was sufficient to reduce MMR by at least 30%. This pathway is mostly found in Level III health facilities, as well as a few Level IVs. This pathway emerges on facilities that have significantly high increases in facility based births (on average, more than 100% from the previous year).

- ***Pathway 4: The Kibaale/Ambulance Pathway, combining Ambulances & VHTs & Medicines & Staff***
  Combining an ambulance transport option with a VHT program, and including the ongoing provision of essential medicines and hiring additional critical staff is also sufficient to reduce MMR by at least 30%. This pathway occurred mostly in Kibaale, where no Boda Vouchers were offered. Although this looks similar to Pathway 2, this pathway cuts across both Level III and IV government facilities, and is more prevalent in facilities that witness lower increases in demand (with less than 100% increase in demand).

- ***Pathway 5: The Private Facility Pathway, combining Delivery Vouchers & Boda Vouchers & Medicines (and assumes VHTs are present)***
  Combining Boda transport vouchers (which assumes VHTs are present) and vouchers for deliveries at private facilities with the provision of essential medicines was the final sufficient pathway to reducing MMR by at least 30%. This pathway only occurred at private health facilities.

We also found that:

- The ongoing provision of essential medicines played a strong role supporting most pathways to success.
- Training activities were not relevant to achieving a 30% or higher reduction in MMR during Phase 1.
- SMGL was unable to decrease MMR by more than 30% at Fort Portal Regional Referral Hospital or at any of the government Hospitals. Indeed, MMR actually increased at two Hospital facilities.

Essentially, the QCA model found that SMGL did not need to implement every activity at every facility to reduce MMR by 30% or more. Following any of the above pathways would have been sufficient to achieve success in their corresponding contexts, and additional activities were not needed. These pathways should be considered when defining the essential package of Phase 2.

Cost Model Findings

We built a cost model that estimates the costs of all births in Uganda, broken down by district, and the incremental add-on costs for SMGL, broken down by activities. Based upon the assumptions of the model, we found that the average cost per GoU-only birth (not including set up costs) is $75.94 per birth across the four districts. These costs range from $17.63 for normal delivery at a government Level III, and up to $176 for births with complications. These costs should be associated with the much higher MMR prior to SMGL Phase 1.

To reduce MMR under Phase 1's full package, the costs for *improving a birth under SMGL* added an average of $136.21, bringing the total to $212.15 and increasing the cost/birth by 179% in the first year. These costs ranged from $153.85 for a normal delivery to $312.21 for those with complications. Broken down by activity, the model shows the incremental cost add-on for each activity across the four districts, broken down by year 1 and year 2 costs (the model costs out up to 5 years):

Activity Break Down of Added Costs Per Improved Birth		
Activity	**Year 1 Cost (w/set up)**	**Year 2 Cost (operational)**
SMGL Management	$18.14	$17.98
Awareness Campaigns	$4.46	$4.17
Village Health Teams (VHTs)	$6.45	$2.80
Boda Transport Vouchers	$3.15	$2.67
Ambulances	$12.96	$1.02
Staff Training	$4.63	$4.40
Staff Mentoring	$2.53	$2.40
Hiring additional personnel	$9.09	$9.07
Private Facility Delivery Vouchers	$78.57	$78.42
Provision of Essential Medicines	$16.13	$17.26
Provision of Essential Equipment	$15.86	$2.46
Upgrades, Renovations and Builds	$30.23	$8.58

System Strengthening	$2.35	$2.20
*Expressed in Cost/Improved Birth, which should be added to GoU Base. Amounts in USD		

Cost Effectiveness Results

Completing Objective 3 involved bringing together the findings associated with the previous two objectives to conduct the cost effectiveness analysis. The following 2x2 cost effectiveness metric presents the findings:

Cost Effectiveness Analysis of Various SMGL Pathways to Reduced MMR Expressed in $/Improved Birth for Year 1

*Note: Amounts do not include $18.14/improved birth for Program Management Costs

	More Expensive Pathways	**Less Expensive Pathways**
More Effective Pathways	Pathway 3 - High Demand Pathway: Ambulance & Medicines & Equipment & Staff = $54.04 Pathway 4 - Kibaale/Ambulance Pathway: Ambulance & VHTs & Medicines & Staff = $44.63	Pathway 1 - Level III Pathway: Boda Vouchers & Mentoring & Medicines &VHTs = $28.26
Less Effective Pathways	*Uncommon Pathway* Pathway 2 - Infrastructure Pathway: Infrastructure & Boda Vouchers & Mentoring & Medicines & VHTs = $58.49 *Not very effective* Pathway 5 - Private Facility Pathway: Delivery Vouchers & Boda Vouchers & Medicines & VHTs = $104.30	*Not relevant or effective* Training: $4.63 (not relevant) Hospitals (not effective)

The costs shown above are expressed in $ of each combined activity pathway/improved birth for Year One and include start-up costs. Each pathway should be understood as the add-on costs that are added to the GoU base costs/birth. These costs are also average costs across the various delivery types, i.e. they average normal births and those with complications. They do not include additional program management costs, i.e. $18.14/birth (i.e. add the program management cost to each pathway total).

Based upon the above, it is clear that Phase 2 mostly likely does not need to implement every activity at every facility to successfully reduce MMR by at least 30%. There are two ways to interpret the above CEA results: simple and complex. Under the simple interpretation, the most cost effective pathway forward would be Pathway 1, a VHT/Boda package combined with Mentoring and Medicines, adding $28.26 per birth plus $18.14/birth for program management costs, $46.40/improved birth. Thus, a more cost effective SMGL would reduce add on costs per birth from $136.21 to $46.40 in Year 1, i.e. around one third of the cost of the full blown package, or a reduction of around 66% per improved birth from pilot year costs.

The complex interpretation recognizes how the effectiveness of each program pathway needs to be understood in terms of the context and level of facility within which it travels (as discussed in more detail in the Objective 1 Chapter). For example, the more expensive Pathway 3 may also be considered cost effective if catchment areas witness at least a doubling of facility-based births from the year prior (which occurred around 25% of the time in Phase 1). Alternatively, should Phase 3 decide not to fund Boda vouchers (with or without Every Mother Counts), as in Kibaale, Pathway 4 would be the recommended choice at Level III facilities over Pathway 1.

Moreover, should Phase 2 expand to Level IV facilities, Pathway 2 is a cost effective approach to supporting Level IV facilities that require infrastructure upgrades. Regardless of the pathway chosen, the provision of essential medicines should be part of any cost effective package.

The analysis above also suggests that supporting private facilities is not cost effective. The analysis also suggests that training – in its current form – is not needed for Phase 2, and implementers may wish to consider revising the training, or shifting those resources to support more mentoring and staff hiring activities. Finally, the lack of success in Hospitals (also noted in the Objective 1 chapter) may suggest that SMGL's design is less appropriate for addressing the challenges facing those facilities.

In terms of the applicability of these findings to Phase 2, we must point out that they are context specific. Different contexts will always imply different pathways. Yet if the problems to be addressed and their corresponding conditions are similar to the challenges addressed in Phase 1, we do expect similar results in those areas.

Recommendations

Based upon the above, we offer two sets of recommendations regarding program design and program M&E:

Program Design Recommendations

Recommendation 1: Instead of administering all Phase 1 activities to each health facility in Phase 2, SMGL stakeholders may wish to consider following one, or a combination, of the following sets of activities:

- At Level III Health Facilities, esp. hard-to-reach rural ones where new staff turnover is high, offer a VHT/Boda package combined with the mentoring of current staff and the provision of essential medicines;
- Should the program be unable to provide Boda vouchers, offer an alternative package based upon ambulances, the provision of essential medicines, VHTs and hire additional staff to support current ones
- Should Phase 2 expand to Level IV facilities, offer infrastructure upgrades to those facilities that need it, mentor current staff, offer a VHT/Boda package and the provision of essential medicines;
- Should implementers anticipate significantly high increases to facility-based demand (above 50%) in both Level IIIs and IVs, offer an improved and fully-supported ambulance referral system (including the ambulances), hire additional staff, and include the provision of essential medicines and equipment;
- Moreover, also consider avoiding Hospitals in Phase 2, or designing a different type of intervention specifically tailored to those unique challenges that SMGL is not always able to address;
- Similarly, consider also avoiding private facilities, as they are the least cost effective approach to reducing MMR (due to the high cost of private delivery vouchers) and less effective in general;
- Consider revising the approach to training, or shifting human resource supports more to mentoring and/or hiring of critical staff; and
- Crucially, ensure that Phase 2 implementing partners have a heightened understanding of the context in order to identify and adapt which pathway is most cost effective for that context.

Recommendation 2: Consider embedding a similar type of cost effective approach into the ongoing implementation of Phase 2 as a means of periodic review, feedback looping, and continuous improvement.

Recommendation 3: Consider conducting a sister CEA exercise in Zambia to better understand the relative success of each program in general and how far the findings generated here may or may not travel to other countries. This may also be useful for understanding what kind of essential package might work in Nigeria.

M&E Recommendations based on conducting the CEA

These recommendations are, in addition to the main body of the report, also based upon the discussion in Annexes 2 and 3.

Recommendation 1: In Phase 2, facility names, areas, complications, medicine names, etc. should be standardized according to a consistent protocol. In regards to facility naming protocols, DHMIS names and spacing should be adopted to facilitate easier table merging with other datasets, although program staff should also identify where DMHIS facility names are inaccurate (mostly in regard to misspecification of facility levels). The naming of complications should also follow a standardized format. These standards should be applied to both partner-level activity data as well as survey/assessment datasets, and, if possible, implementing partners should avoid data entry procedures that require staff to hard code these variables into their database.

Recommendation 2: When adopting the facility as a unit of analysis, designate a staff member to consolidate and maintain an overarching truth table. The truth table should identify what activities and resources are delivered to each facility, when and the corresponding amounts (including currency values), for all partners supporting each facility connected to the program (which suggests that a USG staff member may be the most appropriate choice). Doing so requires that implementing partners streamline relevant program information across various departments (e.g. procurement departments will share the cost of goods delivered, human resource departments will share staff support and turnover data, infrastructure teams will share cost and construction data, etc.). It is also useful to identify minimum thresholds for each activities, i.e. how much of an activity or resource needs to be provided before it is considered "adequate."

Recommendation 3: When supporting an ambulance referral system, partners should develop a logbook that captures both "referred from" and "referred to" fields.

Recommendation 4: At the beginning of the next phase, implementers should identify health facility catchment areas and an agreed upon definition of what constitutes a catchment area. When conducting verbal autopsy surveys, the program should identify an approach for identifying in which catchment area a community based maternal death can be assigned. Enumerators should be trained in these protocols and assign community based deaths accordingly. Ideally, these catchment areas should be standardized and the dataset should be rigorously cleaned to eliminate user entry error.

Recommendation 5: Also in regards to conducting verbal autopsies, implementers should compare their findings with other databases, such as the DHMIS, to increase confidence in the survey's data and flag wide discrepancies that require further exploration.

Recommendation 6: If health facility assessments are conducted (both baseline and endline), assessment designers should identify a means for determining program attribution, e.g. if levels of adequate equipment improve at endline, identify which equipment SMGL provided to attribute the improvement to the program. These assessments should also identify critical staff level of effort at each facility (e.g. part time, full time, concentrated only in maternity ward or timesharing with other wards, etc.) Assessment questions regarding stockout and equipment levels of adequacy should cover the entire period of performance. Partners should also conduct periodic equipment audits to ensure that the equipment they deliver remains in place and in use. The baseline should be used to construct facility level norms that should guide the level of resources delivered to each facility. These norms should be established before implementation begins.

Background[1]

Saving Mothers, Giving Life (SMGL) is a 5-year initiative designed to rapidly reduce deaths related to pregnancy and childbirth through a coordinated approach that strengthens maternal health services in high-mortality settings.[2] SMGL focuses on reducing the "three delays" that prevent women from accessing maternal health services: delays in seeking services, delays in reaching services and delays in receiving quality care. These interventions are focused primarily on the critical period of labor, delivery and the first 48 hours postpartum, when most maternal deaths and approximately half of newborn deaths occur.

Focusing on four districts in western Uganda (as well as in Zambia), SMGL's Phase 1 took place from July 2012 through June 2013. Altogether, 107 Ugandan health facilities that provided delivery care received support from Baylor College of Medicine Children's Foundation Uganda (hereafter Baylor) in the Kabarole, Kyenjojo and Kamwenge districts, and from Makerere University's Infectious Disease Institute (IDI) in the Kibaale district. In close collaboration with the Government of Uganda (GoU), SMGL Phase 1 supported 24 different evidence-based interventions in the community and in health facilities, including:

- Equipping facilities to provide high quality emergency obstetric and newborn care (EmONC) in a geographically strategic way, enabling women with complications to receive care within two hours.
- Improving supply systems so facilities have the equipment, supplies, commodities and drugs needed to deliver high quality EmONC services.
- Training and mentoring skilled birth attendants to provide quality, respectful delivery services, and stabilize, treat and refer emergency cases if necessary.
- Mobilizing the community, through Village Health Teams (VHTs) and other mechanisms, to generate demand for facility deliveries, antenatal and postpartum care while emphasizing birth planning, and encouraging HIV testing and treatment and uptake of family planning services.
- Strengthening linkages between communities and facilities through a 24/7 integrated communications and transportation system, including the use of subsidized vouchers for motorcycle (Boda) transport[3] and free deliveries at private facilities, to help pregnant women access childbirth facilities in a timely manner.
- Improving data collection systems to record pregnancy outcomes, including complications and maternal deaths, and strengthen host country health management information systems.

SMGL activities included routine monitoring during the first year of the project, including baseline and endline assessments before and after Phase 1.

SMGL's Phase 1 results are impressive. Drawing upon an extensive collection of outcome data, the Centers for Disease Control (CDC) found that the maternal mortality ratio (MMR) per 100,000 births fell by at least 30% over the course of the 12 months. This reduction was driven by women's increased access to and receipt of emergency obstetrics and newborn care (EmONC). The Institutional Perinatal Rate also fell by at least 17% in the same time period. Notably, the proportion of expected deliveries taking place at SMGL-supported facilities increased from 46% to 74%.[4] These and other findings reveal that Phase 1 delivered remarkable results, contributing to a rapid decline in the number of women who die during pregnancy and childbirth in SMGL program areas in Uganda (as well as in Zambia).

---

[1] The background section of this report borrows heavily from the *Making Pregnancy & Childbirth Safer in Uganda and Zambia* 2013 Annual Report, found at www.savingmothersgivinglife.org/our_work/reports/annual_report.aspx

[2] Current partners are the American College of Obstetricians and Gynecologists, Every Mother Counts, Merck for Mothers, the Government of Norway, Project CURE, and the US Government [Centers for Disease Control and Prevention (CDC), US Agency for International Development (USAID), Corps, US Department of Defense, and US Office of the Global AIDS Coordinator (OGAC)]. The governments of Uganda and Zambia were central to the partnership and all activities

[3] The Boda Voucher program was funded separately by the foundation Every Mother Counts, and was implement by Baylor through VHTs.

[4] See the "Results at a Glance" report for a summary, found at www.savingmothersgivinglife.org/our_work/infographic_uganda_zambia.aspx

Rationale for this study

Plans are now underway to move SMGL into Phase 2 by expanding to six new districts in Uganda and 12 new districts in Zambia with a long-term goal of national scale-up to save more women's and newborns' lives.[5] To assist, the SMGL partnership has already commissioned two evaluative studies to identify how Phase 1 lessons learned can inform Phase 2. These are the Columbia University Evaluation and the Futures Group Expenditure Study:

- ***Columbia University Evaluation***: The SMGL Global Secretariat commissioned Columbia University to conduct a strategic implementation evaluation to assess the partnership's functioning, identify best practices and barriers to reducing maternal mortality, and gauge implementation quality. The overall findings supported the conclusion that SMGL gave the entire health system increased credibility. However, the study concluded that there "remain mission critical knowledge gaps. One of these is the content of the *minimum essential SMGL package* required to improve maternal survival. Defining such a package is required to scale up the program in the context of limited resources. Going forward, combinations of promising interventions (active ingredients) customized to country needs should be tested..."[6]
- ***Futures Group Expenditure Study***: The United States Agency for International Development (USAID) commissioned the Futures Group to conduct an expenditure analysis of Phase 1 to better understand the level and type of investments made during implementation. The study documented incremental maternal health expenditures made at the district level, and was useful for comparing capital improvement expenditures with those made in the sister SMGL program in Zambia. However, the study did not generate unit cost data (e.g., the cost of improved delivery for a newborn), and also did not catalogue expenditures made at sub-district or health facility levels. As such, the study suggests "the need for a more comprehensive understanding of the cost per delivery in each district after all the projected investments have been made and service delivery takes full effect."[7]

Although these studies have provided useful information that will inform the design and rollout of Phase 2, two important questions admittedly remain unanswered, i.e. what combinations of SMGL activities (i.e. essential packages) should Phase 2 offer to improve maternal survival, and what are the associated costs?

Scope of Work

QEDGroup,llc (hereafter QED) subcontracted SoCha,llc (hereafter SoCha) through the Uganda Monitoring, Evaluation and Learning Contract[8] (aka The Learning Contract) mechanism to perform a Cost Effectiveness Analysis (CEA) of SMGL Phase 1. The overall aim of the CEA was *to determine which aspects of the Phase 1 SMGL intervention are the most cost effective and scalable in preparation for a Phase 2 rollout.* The CEA's primary objectives were to:

- Identify which combinations of SMGL Phase 1 activities are necessary and/or sufficient to improve maternal survival in Uganda;
- Build a model that can project the approximate costs of implementing SMGL Phase 1 activities elsewhere in Uganda; and

[5] For a discussion on how Phase 2 will build upon the experiences of Phase 1, see the Phase 2 Launch Report, found at www.savingmothersgivinglife.org/doc/SMGL_Phase2LaunchReport.pdf

[6] See p. 2 of the Executive Summary of the Columbia University Evaluation Uganda Brief on the official Saving Mothers, Giving Life website version, found at www.savingmothersgivinglife.org/doc/smgl-policy-uganda_v02.pdf Emphasis added.

[7] See p. 7 of the Executive Summary of the Expenditure Study on the official Saving Mothers, Giving Life website version, found at www.savingmothersgivinglife.org/doc/USAID%20SMGL%20Expenditure%20Study%20Executive%20Summary.PDF

[8] The Ugandan Monitoring, Evaluation and Learning Contract is a 5-year, USAID-funded activity dedicated to providing M&E and "Collaboration, Learning and Adaptation" (CLA) services to USAID/Uganda and its implementing partners.

- Combine the above listed analytical components into a third that compares the relative costs of each combination of activities to determine which are the most cost effective. [9]

To meet these objectives, SoCha combined a "qualitative comparative analysis" (QCA) approach with cost modeling to conduct the Cost-Effectiveness Analysis. Further details about the specific Scope of Work can be found in Annex 1: Scope of Work.

Defining CEA

Before defining CEA, it is important to first clarify what it is not. The Phase 1 CEA was not designed to determine if SMGL was effective; SMGL's effectiveness was already established and widely acknowledged.[10] The CEA should also not be confused with cost-benefit analysis, as the later attempts to identify the value of a human life in monetary terms. Finally, this CEA was not designed to audit SMGL Phase 1 spending and expenditures, which would not be useful for determining costs elsewhere.

This report defines CEA as a form of analysis that compares the costs and outcomes (effects) of two or more courses of action. Cost effectiveness itself is expressed in terms of a cost amount in the numerator and an effect in the denominator, e.g. the cost per improved birth. Yet for this exercise, we distinguish between the "traditional" approach and the QCA approach mentioned above. "Traditional" CEAs of maternal health interventions compare a variety of interventions, such as those conducted in other countries, to identify which is the best course of action to achieve the same objective. This can be expressed in a cost/effectiveness table like the one seen below:

The Traditional Approach to Cost Effectiveness Analysis: Comparing Interventions		
	More Expensive	**Less Expensive**
More Effective	Intervention A: More Effective/More Expensive	Intervention B: More Effective/Less Expensive
Less/Not Effective	Intervention C: Less Effective/More Expensive	Intervention D: Less Effective/Less Expensive

Under the traditional approach, Intervention B – More Effective/Less Expensive – is the preferable choice.

The traditional approach is useful when policy makers are considering a variety of distinct interventions to rollout. However, because the SMGL partnership had already decided to move forward with SMGL in Phase 2, there was no value in conducting a comparative analysis primarily oriented to questioning that decision. As such, this CEA did not adopt the traditional approach.

The QCA Approach

Instead, we advanced a new approach that adapted a QCA-effectiveness model with an activity-based costing model. Unlike traditional CEA comparisons that treat each program as a unified whole, we looked at variations in the way SMGL was implemented across health facilities to identify different "program pathways" to reducing MMR. Put differently, we were not just interested in identifying "how" SMGL reduced MMR, but explored if there were in fact a multitude of "hows." To do this, we relied upon QCA to

[9] Taken from Annex 1: Scope of Work in the QED Purchase Order (UMELC-QCA/CEA-001) issued to SoCha,llc. Note the wording of the three objectives has been modified as they have moved from "illustrative" objectives to more clearly defined ones.

[10] See, for example, p.1 of the 2013 SMGL Annual Report, which contains acknowledgements of SMGL's success from the USAID Administrator, the former US Global AIDS Coordinator, and the Director of CDC. Found at www.savingmothersgivinglife.org/our_work/reports/annual_report.aspx and accessed on September 19, 2014.

identify which combinations of SMGL activities led to pathways of success. Concurrently, we built an activity-based costing model of Phase 1 to identify base and incremental add-on costs of each SMGL activity per birth, and layered in district-level hospital delivery data and community-based birth estimates to build a cost model that could be applied throughout Uganda. Finally, we then used the results of the costing model to determine how much each program pathway would cost, thus identifying the essential package of activities that can be cost effectively rolled out in Phase 2:

The QCA Approach to Cost Effectiveness Analysis: Comparing SMGL Program Paths		
	More Expensive	**Less Expensive**
More Effective	SMGL Path A: More Effective/More Expensive	SMGL Path B: More Effective/Less Expensive
Less/Not Effective	SMGL Path C: Less Effective/More Expensive	SMGL Path D: Less Effective/Less Expensive

The findings of each model are discussed in the following chapters of this report. More detailed discussions of the steps we took and decisions we made are found Annexes 2 and 3.

Level of Effort, Unit of Analysis and Data Collection

The SoCha team was comprised of the CEA/QCA team leader and two Cost Model specialists.[11] The CEA began in Uganda in September 2014 and carried forward until the beginning of November 2014. During this time, the CEA team met with and gave several presentations to the CDC and USAID Teams, the SMGL M&E Working Group, and Baylor and IDI implementation teams in Kampala and across the four districts. SoCha team members first met to present and discuss the model, and held subsequent meetings and discussions with implementing partners over critical modeling and data decisions to ensure that both models reflected Phase 1 implementation as accurately as possible. The USAID/Uganda Economist also accompanied the SoCha team on field visits and facilitated discussions with CDC and the implementing partners. From December to January 2015, the CEA team conducted the data analysis, built the associated models, and gave findings presentations to CDC, USAID/Uganda and USAID/Washington, and Implementing Partner audiences to solicit feedback and share analyses. CDC/Atlanta reviewed the report and hired an independent QCA reviewer to assess the findings. Finally, representatives from USAID/Uganda, USAID/Washington, CDC/Atlanta, the SMGL Leadership Council, SoCha and the QCA reviewer came together in Washington, DC in October 2015 to discuss and review the findings. This report incorporates feedback from all of those meetings.

To apply QCA to SMGL, the SoCha team focused on the heath facility (and corresponding catchment area) as the primary unit of analysis and comparison. This required considerable data collection and cleaning, as implementing partner M&E systems focused more on district level aggregates and distributed facility-based data across several operational departments. The previous evaluative exercises mentioned focused more on district level averages and aggregates due to the challenges of facility level data collection. This meant that the SoCha team had to re-create the Phase 1 intervention by compiling facility-level data from a variety of sources. In doing so, the SoCha team amassed a considerable trove of costing, implementation and outcome data. After signing data confidentiality agreements with CDC and the implementing partners, the SoCha team drew upon the following datasets:

- *Maternal Mortality Verbal Autopsy Dataset*: In response to the issue of how to capture community-based maternal deaths that do not appear in official health facility records, the SMGL team

[11] The cost modeling specialists were provided by Cornerstone Economic Research.

conducted an extended Verbal Autopsy of both facility and community based deaths across both baseline (July 2011-June 2012) and Phase 1 (July 2012-June 2013) years. This dataset was the result of 589 extended interviews of relatives of mothers who died during their pregnancy across both time periods, and the dataset contains more than 800 variables. The Verbal Autopsy data was the main data source used by the CDC to report the overall reductions to MMR across the four districts, and the SoCha team also used the dataset to derive facility/catchment area level MMRs (see QCA section).

- *Health Facility Assessment (HFA) Dataset*: Baseline and endline health facility assessments were conducted by the SMGL team to document change in infrastructure, human resources, supplies/equipment, and services. Baseline assessments were conducted in the months preceding the June 2012 beginning of the Phase 1 year, and endline assessments were carried out one month after the conclusion of the pilot year. One key point about this dataset is that it does not establish SMGL attribution, i.e. the questions do not ask respondents if the SMGL program contributed to the changes from baseline to endline. SoCha used this dataset as a means of measuring adequate levels of critical equipment and medical supplies.

- *Pregnancy Outcome Monitoring Survey (POMS) datasets*: These datasets were constructed by the SMGL M&E teams as a result of monthly tracking of maternal outcomes at SMGL-assisted facilities for both baseline and Phase 1 years. The first dataset focused on monitoring outcomes across 112 Level II and III health facilities. The dataset contains 40 variables that are aggregated according to monthly totals across 49,142 births. The second dataset focused on monitoring individual maternal outcomes across 16 Level IV, Hospitals and Regional Referral health facilities. This dataset contains individual entries for 40,525 deliveries, and captures more than 60 variables regarding various complications. The SoCha team used the POMS datasets to identify various complication ratios and procedures, as well as the cross-reference data from other sources and compile descriptive statistics.

- *DHMIS datasets*: DHMIS is a comprehensive online government database of public health indicators across all health facilities in Uganda. Most information goes as far back as 2011. The SoCha team collected district and facility level data from DHMIS to confirm maternal deaths, facility-based deliveries and other facility level statistics. These data were the adjusted upward to account for community-based births beased upon estimates provided by CDC/Atlanta. The district-level demand data used to populate the costing model also came from DHMIS.

- *Futures Group Expenditure Study Data*: As mentioned above, the Futures Group conducted an expenditure study of Phase 1. The study used spreadsheet templates that were emailed to the implementing partners to provide district level expenditure data. Some of this information contained specific line item expenditures on such things as infrastructure expenditures and staff wages. The SoCha team used this information to help inform the Cost Model.

- *Columbia University Evaluation Data*: As part of the evaluation, Columbia University's Principle Investigator (PI) administered a series of questionnaires to 44 of the SMGL Phase 1 facilities, including a knowledge assessment exam to gauge the effectiveness of the SMGL training and a SMGL awareness survey administered to mothers at the facility. The PI shared health-facility level averages of this data and, as per Internal Review Board protocols, ensured respondent confidentiality. We used this information to cross-check findings, esp. regarding the effectiveness of the trainings.

Additionally, SoCha worked with QED to conduct a survey of 40 SMGL Phase 1 Health facilities to identify staffing, equipment supply and delivery bed norms. Finally, the SoCha team collected an extensive amount of implementing partner budget, cost, procurement and implementation data inform both the QCA and Cost models.

The rest of this report is laid out as follows: The next chapter introduces the QCA findings, offers preliminary interpretations, explains the model, its processes and its corresponding data decisions, and concludes with notes on how the modeling process can improve M&E. The third chapter introduces the cost modeling findings, explains how it can be used, its assumptions and data collection procedures. In the final chapter, we bring together the findings of both models to conduct cost effectiveness analyses, and conclude with corresponding recommendations and comments on applying this approach elsewhere.

Objective 1: Identify which combinations of SMGL Phase 1 activities are necessary and/or sufficient to improve maternal survival in Uganda

Introduction

SMGL Phase 1 sought to reduce MMR by addressing the "three delays" that prevent women from accessing maternal health services: delays in seeking services, delays in reaching services and delays in receiving quality care. The Columbia University *Interim Report* presents how the program hypothesized that "the combination of increased information and transport at the community level coupled with availability of better services at health facilities…will dramatically expand health system utilization and maternal and newborn survival." (p. 23). This combined supply-demand-side strategy defined what was referred to as the "whole of health system" approach. Operationally, the whole of health system approach fell across 24 types of activities that fell across three categories of support:

- *Category 1: Activities that Increased Demand and Access to Private and Government Health Facilities*: SMGL sought to shift demand away from traditional birth attendants by mobilizing Village Health Teams (VHTs) to identify and monitor pregnant mothers, raise awareness, and promote facility-based antenatal and delivery care. SMGL also ran mass media campaigns on safe motherhood. In the Kabarole, Kyenjojo and Kamwenge districts, VHTs offered pregnant mothers heavily subsidized Boda (local motorcycle transport) vouchers so they could more easily access health facilities at the onset of delivery. SMGL implementers also offered pregnant women vouchers for free deliveries at private facilities, which were considerably more expensive than at government facilities. In all four districts, SMGL purchased ambulances and communications equipment to transport pregnant women with emergency obstetric complications to receive EmONC.

- *Category 2: Activities that Absorbed Increased Demand and Improved Quality*: SMGL implementers remained cognizant of the potential overload their awareness-raising and increased-access activities would create on the health care system. Indeed, by the end of Phase 1, SMGL-supported, facility-based delivery rates increased by more than 50%.[13] In response, SMGL implementers hired new doctors, nurses and midwives to provide high quality obstetric care; provided additional training in EmONC services, newborn resuscitation, surgical obstetric care and anesthesia; and buttressed this training with ongoing mentoring at select facilities. Key to this strategy were activities designed to shift demand away from the already-overloaded Regional Referral and Hospitals (Level Vs) to Level IIIs and IVs by building the latters' capacity to deliver BEmONC and CEmONC, respectively. SMGL built or renovated operating theaters, provided other infrastructure improvements and sourced new equipment. Most health facilities also received additional equipment, supplies and critical medicines to meet the increase in demand, prevent stock-outs, and maintain the blood supply.

- *Category 3: Activities that Strengthened Health Care Delivery Systems*: SMGL dedicated considerable resources to a series of all-encompassing activities that focused on health-care systems strengthening. These included capacity building activities like training and mentoring in supply chain and other management principles, providing computer-based medical records and communications equipment, and involving district and other officials throughout the entire implementation of Phase 1. These activities also included M&E capacity building and data collection efforts that resulted in the datasets mentioned in the introductory chapter.

More precisely, SMGL's 24 activities appear in the following table in which increased access to quality services are broken out as two distinct types of activities:

[13] According to DHMIS, facility-based deliveries in 99 SMGL health facilities increased by an average of 53.15%.

ACTIVITIES

Demand

- Train and supervise community groups to promote facility delivery and birth preparedness (Village Health Teams [VHTs], Safe Motherhood Action Groups [SMAGs])
- Identify and engage community influencers in safe motherhood (Change Champions)
- Provide basic newborn/birth supplies to incentivize facility delivery (Mama Kits, Mama Packs)
- Run mass media campaigns on safe motherhood, engage community drama groups

Access

- Build new operating theatres with CEmONC capacity
- Expand/refurbish maternity wards, labs, pharmacies
- Renovate mothers' shelters near hospitals for high risk women
- Provide service delivery vouchers and vouchers for transport to facilities and referral to higher-level private facilities
- Buy emergency obstetric and neonatal care (EmONC) equipment
- Buy ambulances, motorcycles, motorbikes, and communication equipment for transportation and referrals
- Form district-level transport committees to improve referral
- Contract with private providers for EmONC

Quality

- Hire new doctors, nurses, and midwives in collaboration with the Ministry of Health
- Train health workers in EmONC
- Train health workers in newborn resuscitation ("Helping Babies Breathe")
- Train doctors in surgical obstetric care and anesthetic officers in anesthesia
- Provide regular supportive supervision to frontline health workers to maintain and improve skills in obstetrics, newborn care and anesthesia
- Provide training and oversight for maternal death reviews
- Provide essential medicines to health facilities

System strengthening

- Build provincial and district health team capacity with SMGL staff
- Strengthen supply chains through training on procurement and stock management
- Build capacity of facility staff to supervise community health workers
- Provide computer-based medical records (SmartCare)
- Strengthen pharmacy, lab, and blood supply
- Support roll-out of electronic district health information system
- Train health workers in data collection and health information systems

BOTH COUNTRIES ■ UGANDA ■ ZAMBIA

Source: p. 22, SMGL Interim Evaluation Report, Columbia University (2013)

Our goal was to determine if, in fact, all 24 activities are needed to obtain a similar 30% one year reduction in MMR in Phase 2. Yet this 30% reduction represents the total change in MMR across all four districts (four-district totals in maternal death and births were taken as a whole across the baseline and endline). This created a challenge for the CEA, as the design assumed comparisons at the facility-catchment area level, i.e. a much lower level unit of analysis than what was used to derive the overall 30% reduction. To be sure, district averages are useful if we want to look at variations in SMGL's success across districts, but this would not be very useful as there are only four districts to compare the various combinations of at least 10 different activities. As such, the bulk of our level of effort for QCA was spent breaking down these district level aggregates and rebuilding a model focused on the health facility and its corresponding catchment area as the primary unit of analysis.

Working closely with Baylor and IDI, the SoCha team modeled Phase 1's implementation by creating a master "truth table" that identified which sets of activities each facility received as well as their corresponding changes in MMR.[14] We were able to do this for 71 health facilities that received Phase 1 support. We then posed the question: *Did SMGL Phase 1 need to implement all of these activities to achieve at least a 30% reduction in MMR?* We answered this question by applying QCA's minimization process to identify which sets of "active ingredients" defined various program pathways to reducing MMR. Although many of the health facilities we modeled also benefited from other SMGL activities in addition to those found in their respective program pathway, activities that sit outside the pathways were not needed because there was counter factual evidence that demonstrated how a facility could still reduce MMR without them. Thus, the program pathways identified below should be understood as sufficient combinations of activities, backed by evidence that shows how additional activities are not needed. Background on QCA as well as details on the steps we took, data we used, and decisions we made throughout the QCA modeling process are found in Annex 2.

Findings

Using QCA, the SoCha team identified five program pathways (combinations of SMGL activities) to reducing MMR by at least 30% over the course of Phase 1. The first four pathways are specific to government-owned

[14] We did not model in the activities associated with the system-strengthening category, as these were focused at the district level and could not be operationalized at the health facility level.

health facilities, and are identified according to the context in which they best fit. The final pathway is associated with private health facilities (for profit and not-for profit):

- ***Pathway 1: The Level III Pathway, combining Boda Vouchers & Mentoring & Medicines (and assumes VHTs are present)***
  Combining a Boda transport voucher option (which assumes VHTs are present) with mentoring of staff, and including the provision of essential medicine is sufficient to reduce MMR by at least 30%. This pathway runs across Level III health facilities with moderate increases to demand for facility-based births (less than 100%).

- ***Pathway 2: The Infrastructure Pathway, combining Infrastructure & Boda Vouchers & Mentoring & Medicines (and assumes VHTs are present)***
  Pathway 2 is a subset of Pathway 1. In facilities that require new or rehabilitating operating theaters, offering the same combination of the provision of critical medicines, Boda Vouchers (which assumes VHTs are present) and ongoing mentoring of critical staff is sufficient to reduce MMR by at least 30%. This pathway is also effective at Level IV government facilities that require new or upgraded or operating theaters.

- ***Pathway 3: The High Demand Pathway, combining Ambulances & Medicines & Equipment & Staff***
  Combining the ongoing provision of essential medicines with ambulance transport services, the provision of essential medical equipment and additional staff was sufficient to reduce MMR by at least 30%. This pathway is mostly found in Level III health facilities, as well as a few Level IVs. This pathway emerges on facilities that have significantly high increases in facility based births (on average, more than 100% from the previous year).

- ***Pathway 4: The Kibaale/Ambulance Pathway, combining Ambulances & VHTs & Medicines & Staff***
  Combining an ambulance transport option with a VHT program, and including the ongoing provision of essential medicines and hiring additional critical staff is also sufficient to reduce MMR by at least 30%. This pathway occurred mostly in Kibaale, where no Boda Vouchers were offered. Although this looks similar to Pathway 2, this pathway cuts across both Level III and IV government facilities, and is more prevalent in facilities that witness lower increases in demand (with less than 100% increase in demand).

- ***Pathway 5: The Private Facility Pathway, combining Delivery Vouchers & Boda Vouchers & Medicines (and assumes VHTs are present)***
  Combining Boda transport vouchers (which assumes VHTs are present) and vouchers for deliveries at private facilities with the provision of essential medicines was the final sufficient pathway to reducing MMR by at least 30%. This pathway only occurred at private health facilities.

We also found that:

- The ongoing provision of essential medicines played a strong role supporting most pathways to success.
- Training activities were not relevant to achieving a 30% or higher reduction in MMR during Phase 1.
- SMGL was unable to decrease MMR by more than 30% at Fort Portal Regional Referral Hospital or at any of the government Hospitals. Indeed, MMR actually increased at two Hospital facilities.

Essentially, the QCA model found that SMGL did not need to implement every activity at every facility to reduce MMR by at least 30%. Following any of the above pathways would have been sufficient to achieve success in their corresponding contexts, and additional activities were not needed. These pathways should be

considered when defining the essential package of Phase 2, and will be re-visited in terms of their cost effectiveness in the final section of this report.

Explaining the Findings

QCA, in and of itself, does not provide explanations for *why* the above-mentioned pathways were sufficient. In this section, we offer tentative explanations for why these pathways were found. In doing so, we draw a distinction between two types of development hypotheses: SMGL's theoretical hypothesis and SMGL's implementation hypothesis. The theoretical hypothesis refers to aspects of SMGL's theory of change that are firmly rooted in *medical science* and evidence, e.g. anticonvulsants are necessary to respond to certain complications at birth and are therefore useful for reducing MMR. By contrast, the implementation hypothesis refers to aspects of SMGL's theory of change that are rooted in *implementation science* and the specific health systems challenges unfolding in the Ugandan context. For example, although anticonvulsants are essential to addressing certain delivery complications, their regular provision under SMGL was not always required to reduce MMR in certain facilities, because, for example, local health facilities already had access to these medicines or could easily obtain them through informal networks; etc. In what follows, we explain the various program pathways through the lens of implementation science using the three delays framework, i.e. why was SMGL able to reduce these three delays enough to reduce MMR by 30%?

Explaining the implementation hypothesis requires that we not only examine how the program was implemented, but also how it interacted with context and other relevant factors that sat outside of SMGL's control. When using the term "context" below, we are referring to factors that are not directly part of the intervention, but nevertheless appeared to play a relevant role, such as geographic factors (rural, urban), institutional factors (facility type, private vs. government, etc.), as well as organizational factors (high staff turnover rates, high increases in demand, etc.).[15] We drew these context variables from data that was already available, but no doubt additional context variables could add explanatory power to the findings, such as cultural factors, etc. In fact, QCA best practice usually involves several iterations of model building, data collecting and interpretation to more precisely identify which set of factors do the best job of defining "context." In this way, QCA is not just an analytical technique to confirm program effectiveness, but also facilitates an interactive process of dialogue between evidence and data.

Given their similarity, the combination of activities associated with Pathways 1 and 2 are best discussed in tandem with each other. Similarly, Pathways 3 and 4 are ostensibly similar, but in face have some key underlying differences that justify their separation. Finally, Pathway 5 is limited to private facilities, but also need to be discussed in the wider inability of SMGL to reduce MMR in private facility catchment areas relative to government ones. For each Pathway, we discuss each activity and include the percentage of facility-catchment areas covered under SMGL.

Level III and Infrastructure Pathways

- Pathway 1: The Level III Pathway, combining VHTs& Boda Vouchers & Mentoring & Medicines (emerged on 24% of government facility catchment areas)
- Pathway 2: The Infrastructure Pathway, combining Infrastructure & VHTs & Boda Vouchers & Mentoring & Medicines (emerged on 16% of government facility catchment areas)

Pathways 1 & 2 are essentially the same pathway, but slightly differ in terms of where they are found. Pathway 1 emerged exclusively on Level III pathways and covered 24% of all government facility catchment areas modeled, while Pathway 2 added those facilities that also required infrastructure upgrades (several Level IVs and Rwamwanja HC III). As such, we treat infrastructure upgrades as a context variable that should

[15] Interestingly, we did not identify any patterns that could be explained by using implementing partner variables. Although their menu of activities differed, both partners appear to have been more or less equally effective in delivering those activities.

guide program implementers on which appropriate essential package to choose, i.e. follow Pathway 2 at facilities that require renovating/building operating theaters.

Boda Vouchers

Boda vouchers were subsidized transport vouchers in which were sold to pregnant mothers by VHTs at a subsidized price. Expecting mothers could then use the vouchers for Boda (local private motorcycles) rides to the health facility at the onset of labor. Boda drivers would then redeem the vouchers at the facility and received above-market fees. Recall that Boda vouchers were not directly paid for by the USG, but were an add-on activity funded by Every Mother Counts. Boda vouchers were also only administered by Baylor University, and therefore not applicable to the Kibaale district. According to the results of our QCA-based model, SMGL's Boda voucher activities, in combination with those that address supply-side improvements, turn out to be a key element for increasing demand and reducing MMR at Level III facilities. SMGL program data reveals that, when given the opportunity, most pregnant mothers chose to purchase a Boda voucher well in advance of their delivery, and that these vouchers assisted more than 70% of facility-based deliveries across most Level III facilities in Kaborale, Kamwenge and Kyenjojo districts. Clearly, Boda vouchers are an excellent tool for stimulating demand for facility-based births at Level III facilities.

Yet during the course of our investigation, two inter-related questions arose: First, how could our QCA model identify Boda vouchers as a sufficient condition without also identifying VHTs?(the vouchers were sold to mothers by VHTs); and Second, why were Boda vouchers so effective?:

The Boda/VHT Paradox

Pathways 1, 2 and 5 all identify Boda voucher program as constituent elements, which is peculiar considering that VHTs themselves were the main distributors of Boda vouchers, i.e. how can we see one without the other? The answer is found in the way the Boda voucher program was administered by Baylor University. Under their guidance, VHTs were assigned to a facility catchment area to, among other things, encourage mothers to buy Boda vouchers to travel to a health facility to give birth. However, mothers who purchased the vouchers were by no means constrained to give birth at the facility in which the VHT was employed, and were instead free to instruct the Boda driver to take her to an alternative facility. Thus, although the VHT component was a necessary precondition for any mother to have access to a Boda voucher within Baylor's three districts, not all of the facilities found within those three districts needed to employ a VHT to benefit from the Boda voucher program and could still receive mothers who purchased vouchers from VHTs assigned elsewhere.

In fact, 15% of the health facilities we modeled hosted mothers who redeemed Boda vouchers but did not have any assigned VHTs to them. Yet this only partially explains the paradox, as Pathways 1, 2 and 5 cover a much higher number of facilities than those captured in the 15%. The remaining portion of the explanation is found in understanding how QCA works. Recall that QCA rests upon principles of logical minimization, e.g. when comparing facility A, which received, say, Boda Vouchers & Medicines & VHTs, with facility B, which received, say, Boda Vouchers & Medicines (but not VHTs), QCA logically concludes that VHTs cannot be part of the cause that reduced MMR, and VHTs drop out of both facilities. QCA then applied this finding to all the other facilities similar to facility A; logically concluding that VHTs were irrelevant. As such, we are obliged to point out that Pathways 1, 2 and 5 all *assume* that a VHT component is still in place, even though many Level IIIs can benefit from the Boda voucher program without employing their own VHTs.

Why were Boda Vouchers so effective?

When understanding why Boda Vouchers were so effective, it is necessary to understand the entire incentive structure driving VHTs, Mothers and Boda drivers. The Boda Voucher program actually incentivized all three members of the delivery transport chain: First, VHTs purchased the vouchers directly from Baylor at around 500 UGS per voucher ($.20) and then sold them to the mother for around 1000 UGS ($.40), i.e. a

100% markup that was the VHTs to keep. Mothers in turn were incentivized to buy these vouchers, as the normal market value of a Boda ride to deliver at a hospital averaged around 3000 UGS ($1.20), thus giving the mother a discount of around 60% and saving her around 2000 UGS ($.80). Finally, Boda drivers would receive around 10,000 UGS ($4.00) per voucher once they carried the mother to the delivery area, i.e. a markup of around 330%.[16] The Boda voucher program thus not only supplemented VHT allowances with a commission-based income, but also passed on cost savings to the mother and rewarded Boda drivers with above market rates.

The Boda voucher incentive structure effectively addressed the first two delays in the 3-delays framework, i.e. reducing delays to seeking care and delays accessing that care. The first delay was reduced when a mother purchased a Boda transport voucher, as she had already decided to deliver in a facility. The second delay was reduced because the Boda driver received above market rates for transporting the mother, which encouraged him/her to give preferential treatment to the mother, even at the expense of other clients who pay the market rate (as well as estimates that 10,000 UGS is the average daily wage for Boda drivers in the area). This preferential treatment given by Boda drivers also helped reduce transport times to the facility relative to other common transport modes (walking, taking a bus, waiting for a Boda to be free, relying upon friend/family for support, etc.). Finally, VHTs were incentivized to sell as many Boda vouchers as possible; leading some to conclude that pregnant mothers were pressured by VHTs to buy these vouchers and deliver at health facilities. We were not able to investigate this claim, but exit interview survey data provided by Columbia University shows that around 17% of respondents they surveyed (mothers who had just given birth at a facility) did feel pressured to deliver in a health facility.[17] This suggests that coercion may have been contributing factor to voucher sales, but not the defining.

The success of the Boda voucher activity holds implications for the commonplace view that many rural Ugandan women prefer traditional birth attendants over facility-based births. Leaving aside the small portion of mothers who were reportedly coerced, the high demand for subsidized Boda vouchers does suggest that most pregnant mothers already understood the value of facility-based births, but couldn't afford the associated transport costs. These barriers can be reduced with seemingly small subsidies to the market price.

Mentoring

The mentoring of current essential staff, combined with other activities, played a crucial role in reducing MMR on Pathways 1 and 2, as it helped build the capacity of staff to deal with the large increase of facility based births brought on by SMGL's demand side activities (e.g. VHTs and Bodas). Note that across all pathways, our model found that either staffing or mentoring, in combination with other activities, were sufficient to reduce MMR, but in no instance were both supports needed. But why would mentoring work in some contexts while staffing work in others? Our investigation into this question revealed that most Level IIIs and IVs that travelled along Pathways 1 and 2 were in more remote, rural areas. Some of these did initially benefit from additional staff members, but most of those staff were younger midwives coming out of university. Many of these staff resigned or transferred shortly after their arrival. As one midwife explained to us, new midwives were harder to retain because they were not prepared for or interested in longer term postings in very remote areas, and were less interested in long-term village life. By contrast, mentoring in midwives who had already made the commitment to live in these areas was a safer investment.

Why was Training Irrelevant?

When interpreting why QCA identified mentoring as part of separate combinations of activities that are sufficient to reduce MMR, it is also useful to point out the irrelevance of training. By irrelevance of training, we are not suggesting that training was completely irrelevant, nor that it didn't support any positive outcomes. Instead, we are suggesting that it is not necessary to include training in the essential package if the goal is to

[16] Note mothers would only receive transport to the facility. They received no assistance returning home.
[17] This was in response to the question: Did you feel pressured to deliver at a health facility for this baby?

reduce MMR by 30%. Recall that none of the identified pathways listed training of essential staff as a part of a combination of activities sufficient for success. During our field visits, some attendees we interviewed confessed that the trainings were mostly "refreshers," while a technical advisor to the training activity also commented that the refresher trainings lacked potency. This suggests two possible explanations: 1. Refresher training is not an appropriate type of activity to address this type of social problem (persistently high levels of MMR); or 2. The quality of the implementation of the training was low (i.e. it may be better to find another implementer). In fact, both training and mentoring activities were implemented by the same organization in all four districts (the Association of Obstetricians and Gynecologists of Uganda). In contrast to the training, mentees described their mentoring sessions as more meaningful because mentors directly observed their behavior and explained their suggested corrections in applied terms. As such, we suspect that the quality of training provided was not the issue, but instead that the refresher content was not the right fit for this intervention.

The Provision of Essential Medicines

The provision of essential medicines[18] was another crucial activity that not only defined Pathways 1 and 2, but also emerged as a constituent member of all the pathways identified above. Given the substantial increases in demand for facility-based care as a result of SMGL, this should come as no surprise and confirm that most health facilities in the area needed substantial assistance addressing these changes. However, the reasons for why this is so varies according to context. The need for the provision of essential medicines on Pathways 1 was done in combination with significant increases in demand created by the VHT and Boda voucher activities at Level III facilities. Unlike the higher level facilities, government Level IIIs receive essential medicines and supplies from the National Medical Stores (NMS) through an "informed push system" of quarterly cycle deliveries set to a fixed amount derived on an annual basis (known as the Essential Medicines Kit (EMK) revisions). These amounts remain fixed throughout the year, regardless of subsequent variations in patient volumes. For many Level IIIs, EMK revisions took place prior to the onset of VHT- and Boda-driven increases in demand, and thus the provision of essential medicines most likely made up for unexpected shortfalls.

The provision of essential medicines also formed in Level IVs that ran across Pathways 2, 3 and 4, where different supply chain systems were in place. In the higher level government health facilities (Level IV and above), essential medicines and supplies were delivered by NMS according to a "pull system," in which health facilities submit reply requests on a quarterly basis based upon fluctuations in actual demand. Although SMGL provided additional supply management training to these higher level facilities so they would be better able to respond to the expected increase in demand, our examination of partner supply requisition records revealed that SMGL received substantial requests from Level IVs and above for additional medicines up until the end of Phase One, i.e. they still struggled to meet the increases in demand.

Finally, the provision of essential medicines also emerged as part of Pathway 5 in regards to private health facilities. Private health facilities typically purchased supplies and medicines through local providers as well as the Joint Medical Stores (JMS), a wholesale not-for-profit enterprise originally set up to support faith-based medical care service providers. Respondents at private health facilities indicated that JMS and its affiliated network was a more reliable supplier than the NMS, due partly to its pseudo market orientation as a not for profit that operates just above cost. However, given the opportunity for cost savings, SMGL-supported private health facilities also benefited from the provision of essential medicines.[19]

---

[18] By essential medicines, we are referring to regular supplies of antibiotics (including Ampicillin, Gentamicin, Metronidazole, and Procaine benzyl penicillin G), Magnesium Sulfate, Oxytocin, Atropine, and Ketamine. Although there were other medicines that could've made this list, we were able to identify stockout levels for these five for the entire Phase 1 as a result of the SMGL Health Facility Assessments.

[19] In the 18 private health facilities we modeled, Midas Touch was the only heath facility that was able to sufficiently avoid stock-outs. For the most part, Midas Touch was supported by Baylor, but we were informed that the provision of essential medicines fell to the STRIDES program.

Essential medicine stockouts have long been a problem in Uganda, yet in some ways SMGL would have exacerbated this problem had Phase 1 not supplied additional medicines. Across all levels of health facilities, representatives mentioned how they oftentimes were still able to draw upon informal networks to fill gaps during periodic stockouts; yet only when a handful of individual facilities in the area faced shortages. By contrast, these networks would fail to deliver when multiple facilities in the area faced stockouts at the same time, resulting in a widespread stockout shock. Because both SMGL implementers maintained ongoing stock in their field-based warehouses, they were more able to sustain a wider buffer zone of backup medicines that helped address the challenge of reducing MMR even during times of systemic shortage. As such, we conclude that the provision of essential medicines should be strongly considered as a constituent element of Phase 2, because doing so addresses different types of supply chain problems with different needs.

Infrastructure

Pathway 2 is unique to health facilities that received new or infrastructure upgrades to their operating theaters and/or maternity wards. Facility upgrades enabled many Level IV facilities to provide CEmONC services and relieve some of the emergency obstetric care burden from the hospitals, especially in regards to caesarian sections.[20] Indeed, Columbia University estimated that these upgrades tripled the capacity of the four districts to address obstetric complications. In this Pathway, the mentoring of Medical and Anesthetic Officers was always found, suggesting that their focus on addressing obstetric complications was a welcome human resource complement to the enhanced physical capacity of these as a result of the upgrades.

High Demand and Kibaale/Ambulance Pathways

- Pathway 3: The High Demand Pathway, combining Ambulances & Medicines & Equipment & Staff (emerged on 31% of government facility catchment areas)
- Pathway 4: The Kibaale/Ambulance Pathway, combining Ambulances & VHTs & Medicines & Staff (emerged on 16% of government facility catchment areas)

Pathways 3 and 4 are ostensibly similar, as both combine Ambulances with Medicines and extra Staffing, and differ only in terms of additionally providing either equipment or VHTs. Yet they do retain a few essential differences that justify their separation. Pathway 4 was only found in Kibaale district, where Boda vouchers were not offered, and where VHTs faced a difference incentive scheme. Pathway 3, which does contain a few Kibaale-based facility catchments, otherwise almost exclusively applied to health facility catchment areas that witnessed extremely high increases to facility-based births. These increases placed considerable strains on facility resources than in most other SMGL-supported facilities. Below we first discuss the role of VHT and Ambulance demand-side activities in contrast to Boda vouchers, and then turn to supply side contributions of medical equipment and extra staffing.

Ambulances vs. Bodas

The health facilities that travelled Pathways 3 and 4 demonstrate that SMGL's efforts to build and sustain a functioning ambulance transport and referral system can be effective even if Boda voucher package is absent. As part of the ambulance package, SMGL implementers worked to strengthen the referral system both at the facility and district health office levels. SMGL purchased various types of ambulances to exclusively transport mothers with obstetric emergencies; most of which involved caesarian sections. According to the POMS datasets we received, around 24% of all deliveries that took place in SMGL-supported Level IV and above health facilities involved caesarian sections (representing 9% of all deliveries when we expand the denominator to include deliveries at SMGL-supported Level IIIs).[21] Thus, it is no surprise to find that

[20] Columbia University estimated that these upgrades tripled the capacity of the four districts to address obstetric complications. See the Final Report, p. 53.

[21] By contrast, around .65% of all deliveries, or 1.7% of those at Level IV and above, involved antepartum/postpartum hemorrhages

ambulances played a decisive role in high delivery rate contexts and in Kibaale where no Boda vouchers were present, as obstetric ambulances were particularly useful for shifting mothers with high risk complications away from Level IIIs and some Level IVs to receive more specialized care elsewhere.

During our discussions, some implementing partners emphasized the *synergistic* nature of the Boda voucher program and the ambulance referral system, i.e. that the benefits of Bodas bringing the broader population of mothers with normal deliveries faster to Level IIIs should combine with the benefits of ambulances prioritizing high risk mothers with obstetric emergencies to Level IVs and above. We ran several iterations of the QCA model to test this proposition but never encountered a situation in which the provision of BOTH ambulance and Boda vouchers were members of the same pathway. However, it should be noted that we were only able to model the facilities that had patients referred to it and not the percentage of mothers who were referred from other facilities.[22]

VHTs vs Bodas

VHTs were designed to increase demand for facility based care through monitoring and awareness raising, help foster the early detection of obstetric complications by promoting antenatal care visits, and assist mothers with locating the appropriate facility in their catchment area. VHTs address the first two delays a mother faces when giving birth, i.e. delays to deciding to seek facility-based care, and to reaching the facility. In contrast to the Baylor VHT model, VHTs in Kibaale did not offer Boda vouchers, and most mothers had to find their own means of transport to the facility (unless they faced complications and could rely upon an ambulance). Also in contrast to the Baylor VHT model, IDI invested more into training VHTs and provided more material incentives to its VHTs (e.g. higher per diems, more air time, bicycles, etc.). IDI staff hypothesized that their model was potentially more effective, as the additional energy associated with establishing a voucher program, convincing and recruiting Boda drivers, selling the vouchers to the mothers, etc. (i.e. the Baylor model) would constrain VHTs to reach fewer mothers.

Using the program date we collected and the results of each model, we were able to test this hypothesis and see which model could be associated with higher performance. We compared the average delivery rate increases to understand which model did a better job of convincing mothers to give birth in facilities, as well as average reductions to MMR to see which model, when combined with other activities, could be associated with higher reductions. The results are:

Comparison of IDI and Baylor VHT Model		
	Avg. Delivery Rate Increase[23]	Avg. MMR Decrease
IDI VHT Model	69%	78%
Baylor VHT Model	97%	79%

The data suggests that while both VHT models can be associated with similar levels of MMR reduction, the Baylor (Boda voucher) model was overall more effective in convincing mothers to deliver in facilities. While it may be true that the Boda voucher program required additional time and resources to establish and maintain, the data suggests that the returns are worth the extra effort, as mothers are more able/willing to have facility based births if they have access to subsidized transport. Nevertheless, the results associated with the IDI VHT model remain impressive and should be considered in situations where a Boda voucher program is not feasible.

[22] We were unable to model referred from data because IDI did not capture this information until the end of Phase 1 and thus our dataset was incomplete. A side lesson learned should be that all future partners who support ambulance transport and referral activities should be required to capture both referred to and referred from data.

[23] Taken from partner VHT and Boda voucher data, and combined with DHMIS data over the base and SMGL Phase 1 years.

Staffing

Both Pathways 3 and 4 include staffing as a constituent part of each essential package. In contrast to mentoring, staffing provided new staff, mostly midwives, to address human resource shortages and fill critical gaps. During our field visits, midwives at Level IIIs did mention how the SMGL-sponsored increase in facility-based deliveries considerably intensified their workload (and fueled burnout), so the additional critical staff were very well received.[24] Thus, the simple explanation is that hiring more staff can been seen as a more potent support than just mentoring current staff alone. This becomes especially clear on Pathway 3, where extreme increases to facility based births overwhelmed current staff and demanded levels of effort that were unsustainable over time. The choice between mentoring and staffing likely depends on facility and individual factors such as the ability of each facility to efficiently absorb new staff into their current division of labor (a strength suggested by travelling Pathways 3 and 4), or the ability of current staff to benefit from mentoring activities in rural and lower demand contexts (as suggested by Pathways 1 and 2). As such, each different pathway's success should be viewed as context specific to the human resource and organizational challenges their corresponding facilities faced.

Equipment

In addition to the provision of ambulance referral services, Pathway 3 also contains the provision of essential medical equipment as another active ingredient in the combination of activities sufficient to reducing MMR. By this, we are referring to adequate equipment that is present in the labor ward for emergency obstetric care.[25] It should be relatively intuitive that the provision of essential medical equipment would play a strong role in reducing MMR in the four pilot districts, especially in the context of significant increases in facility-based deliveries. Conversely, in most facilities that had below average increases in facility-based deliveries, additional equipment may not have been needed. Yet we would also like to point out that SMGL's success with ensuring that labor wards had adequate equipment for emergency obstetric care was mixed. When asked if the following was consistently available and functioning, around 85% of respondents indicated that new born resuscitation equipment and infusion sets were adequate; 49% and 41% indicated that maternal resuscitation equipment and fully stocked emergency trolleys were adequate, respectively, and 29% and 23% indicated that vacuum aspiration and vacuum extraction equipment were adequate.[26] Implementing partners at Baylor offered that one potential reason for the disconnect was due to the fact that they did not receive the results of the baseline Health Facility Assessment – which identified baseline levels of adequacy for obstetric care equipment in the labor wards – until months after the project and procurements had begun. For Phase 2, SMGL implementers should consider improving the timing of baseline results so that it informs the procurement and allocation process sooner rather than later.

Private Facility Pathway

- Pathway 5: The Private Facility Pathway, combining Delivery Vouchers & Boda Vouchers & Medicines & VHTs (emerged on 63% of private facility catchment areas)

Pathway 5 combines Boda transport vouchers (which assumes VHTs are present) with vouchers for deliveries at private facilities and the provision of essential medicines as an approach sufficient to reducing MMR by at least 30%. Pathway 5 is exclusive to private facilities. In the model, it applied to

[24] Columbia University also made similar observations, see especially p. 47 in their March 2013 Interim Report.

[25] Specifically, we refer to equipment to perform vacuum extraction, manual vacuum aspiration, maternal and new born resuscitation are consistently available and functioning, as well as the presence of full infusion sets and fully stocked emergency trollies.

[26] These averages were taken from the Health Facility Assessment endline survey. According to the equipment delivery data we received from the implementing partners, the top seven types of equipment SMGL provided to health facilities (in order of most common to less common) were delivery sets, dispensing trays, drip stands, autoclaves, caesarian sets, suction pumps and oxygen concentrators.

close to two-thirds of the successful facilities under SMGL. The remaining pathways identified by QCA were either not applicable as recommendations or were specific only one facility (see Annex 2 for details). Pathway 5 applies to both for profit and not for profit (typically faith-based). Across all four districts, SMGL supported private health facilities with fewer activities, but also subsidized delivery expenses through an additional voucher program. Mothers could deliver for free at these facilities, and the vouchers covered both normal deliveries and those with complications. Combined with Boda vouchers to improve access, the private facilities we modeled witnessed a below average increase in the delivery rate of 36.8% in Phase 1.

During our field visits, SMGL implementers informed us that private facilities are generally viewed by the populace as higher quality service providers relative to government facilities, but delivery costs were prohibitive for most mothers so attendance was lower. While this may be the case, 38% of the SMGL-supported facilities we modeled did not reduce MMR by at least 30%. Moreover, one-third actually witnessed increases to MMR during Phase 1, with an overall average increase to MMR by 44% across these facilities. These bifurcated outcomes suggest that certain types of private facilities may not have been able to meet the increases in demand created by transport and delivery vouchers, even though they were relatively lower than those in government facilities.

We do not have a strong explanation for SMGL's limited success at private facilities, and were not able to identify any commonality for successful/unsuccessful facilities according to orientation (private not for profit vs. profit), level or district area. However, we suspect that the bifurcation may have to do with the fact that the bulk of support took the form of cash subsidies via the delivery voucher program instead of direct support. Although we were unable to confirm, it could be that private facilities did not reinvest enough of the revenue produced by the voucher program back into the facility. Indeed, respondents at two private facilities we visited indicated that, if given the choice, they preferred to receive essential medicines instead of additional staff members. The reason why was because they believed that their staff were strongly committed to their patients, and could be asked to work extra hours in the face of higher demand at no additional costs. In fact, their hypothesis is supported by the findings of Pathway 5, in which almost two thirds of the private facilities we modeled did significantly reduce MMR. Yet for the other third, intensive, prolonged increases to delivery rates at these facilities most likely cannot be sustained without additional investments.

Hospitals

We were able to model five of the six hospitals, including Fort Portal Regional Referral, in the Phase 1 pilot.[27] Out of the five, only Virikia Hospital (a private, non for profit in Kabarole) met the 30% threshold with a 39% decline in MMR. Kagadi and Kyenjojo Hospitals witnessed 22% and 16% declines to MMR, respectively, while Kabarole and Fort Portal Regional Referral experienced increases to MMR. Moreover, the average number of facility-based births actually declined by 4% for these five hospitals. Although this declining rate should be viewed as a success in terms of SMGLs goal of redistributing facility-based births to Level IVs and Level IIIs, the decline does not help explain why Hospitals were mostly unsuccessful in terms of reducing MMRs to comparable levels.

As in the case of private facilities, we can only speculate on the reasons, but we do suspect that the main reason is connected to the fact that the processes and systems in place at hospitals are qualitatively different (and go beyond volume) than those we find in Level IIIs and IVs. As such, it may be that the types of solutions offered by SMGL – equipment and medicine top ups, refresher training, etc. – were not potent enough to address the challenges facing hospitals. Moreover, we were informed that populace generally preferred to travel the extra distance to a hospital, even if an adequate Level III or IV is close by, as they perceived the larger as having more resources and better able to deal with complications. But the added delays to reaching the facility may have factored into the higher rates (although this should've been a consistent factor running across both baseline and SMGL periods). Finally, an additional reason may have to

[27] Maternal death rate information was not found for Kidal Hospital (or there were no deaths), a private, not-for-profit hospital in Kabarole.

do with the ambulance referral system. Although there were noticeable improvements to the referral system, one of the consequences was that most of the very high-risk obstetric emergency burden was shifted to the Hospitals. Indeed, implementing partners and CDC Uganda representatives informed us that because of the referral system improvements, Level IV facilities were able to identify a number of highly-specialized complications and shift these patients to Fort Portal and the Hospitals.[28]

No doubt a variety of factors contributed to the mixed performance of Hospitals under Phase 1 that require further investigation. Regardless, these findings do suggest that Phase 2 implementing partners strive to ensure that maternal delivery gains made at one level do not translate into losses at another level. Consequently, overly-emphasizing the effectiveness of only one Pathway without understanding its relationship to the rest of the health care system is problematic. As such, we reiterate the need to ensure that each Pathway is understood in terms of an associated context, and add that understanding how high risk factors may be unintentionally shifted to another context can be just as important.

Conclusion

The QCA model found that SMGL did not need to implement every activity at every facility to reduce MMR by 30% or more. Following any of the above pathways would have been sufficient to achieve success in their corresponding contexts, and additional activities were not needed. These pathways should be considered when defining the essential package of Phase 2.

[28] Unfortunately, we were unable to fully test this hypothesis, as the ambulance referral data we received from IDI and modeled only contained "referred to" data, but did not contain "referred from" data. Modeling the referred from data would've allowed us to determine to what extent these high risk patients contributed to the observed changes to MMR.

Objective 2: Build a model that can project the approximate costs of implementing SMGL Phase 1 activities elsewhere in Uganda

Introduction

This section is dedicated to discussing the other side of the CEA approach – the cost model. The SMGL Costing Model was designed to fulfill the following objectives:

1. Model the cost of each of the components of the SMGL program as it was implemented in four districts of Uganda during Phase I;
2. Model the cost to the Government of Uganda of the maternal health care system in the four districts where the SMGL program was implemented, so that the incremental cost of the SMGL interventions can be determined; and
3. Build a model that is fully scalable so that scenarios for the rollout of components of the SGML program to other districts in Uganda can be costed.

The cost model should not be confused with the Futures Group *Expenditure Study*. Whereas Futures Group recreated expenses at the district level, we modeled costs at the facility level. Modeling costs involves identifying a. demand; b. inputs needed to meet that demand; and 3. the price of those inputs. In doing so, we didn't seek to historically recreate SMGL Phase 1 expenditures the way Futures Group did, but instead gathered only enough information to create cost ratios that could be applied elsewhere to varying levels of demand. Moreover, because we needed to understand what inputs were required, we had to go in much greater detail at lower levels than what the Futures Group did.

The cost of the activities in the maternal health care system are driven by the demand for services, which drives the quantity of services required, the resources or inputs consumed in the activities and the price of these activities. When costing a program, it is important to understand the flow of activities, the facilities involved in the activities, the range of inputs used and the price of these inputs. Many of Phase 1 activities were also demand driven, for instance the provision of transport vouchers. However, other activities depended on the program design and explicit decisions made by the program implementers to provide a specific level of support, for example the purchase of additional equipment for facilities and the upgrading of infrastructure. In these later instances, assumptions were developed that sought to replicate the kinds of decisions that were made in the SMGL Phase I implementation.

Even though we introduced a great deal of complexity into the model with the process maps, our approach always came back to the simple costing formula: Cost=Demand x Inputs x Price. This was in turn divided into two different conceptual methodologies. Broadly speaking, the two approaches were used:

i. *Activity based costing* for those components that entail implementing a process such as delivering a service or providing support. These processes can be once-off in nature (i.e. set-up activities) or they can be continuous in nature (i.e. service delivery activities).
ii. *Capital costing* for those components that involve building facilities and supplying equipment such as vehicles and medical equipment.

The methodologies used to cost the maternal health care system were:

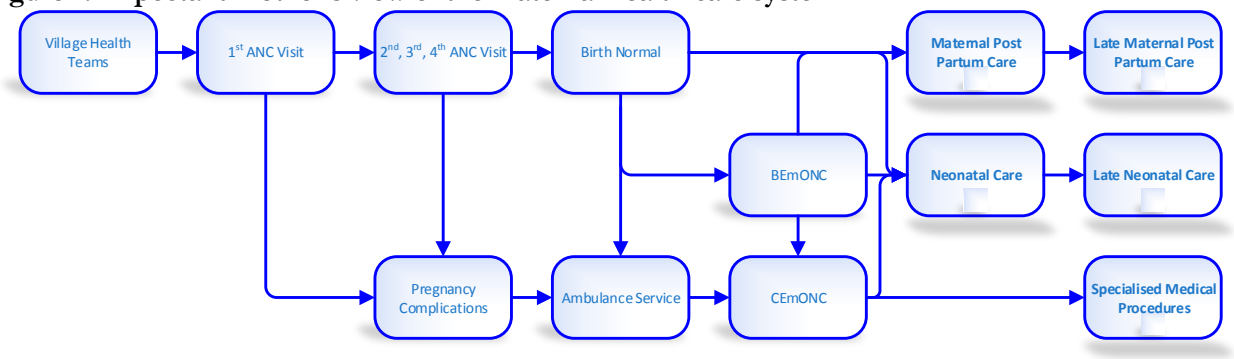
iii. *Institution based costing* for staff and equipment:
- average staffing levels for the different health care facilities are used to calculate the cost of personnel employed at health facilities to deliver maternal health services; and.
- the amortization cost of equipment the health facilities required to deliver maternal health care services, adjusted by a factor that reflects the undersupply of such equipment.

iv. *Cost of procedures* for medicines and supplies used within the maternal health care system.

Design of the SMGL Costing Model

As noted above, the "expectant mother view" to the maternal health care system informs the design of the costing model where appropriate. Understood in broad terms, the main elements of the expectant mother view can be summarized as follows:

Figure 1: Expectant mother's view of the maternal health care system



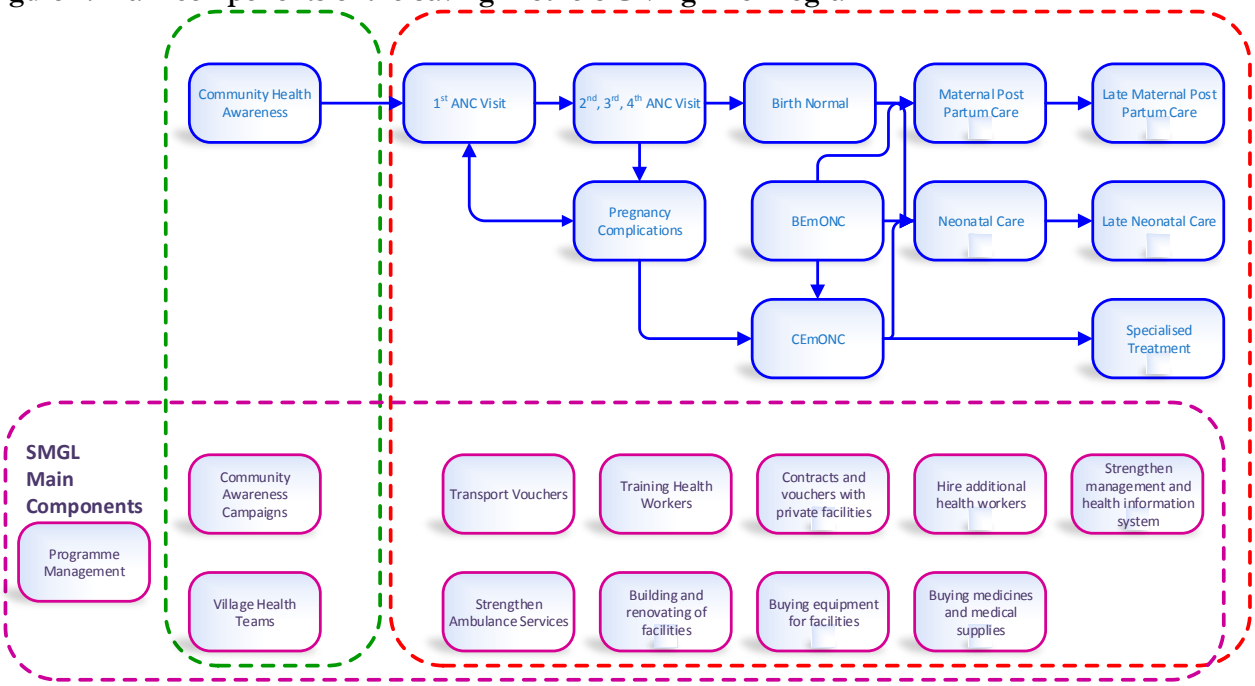
This summary should be understood as broad baseline processes that were already in place prior to SMGL's commencement (i.e. the Ugandan health care system costs). The associated costs of each of these elements are highly influenced by varying levels of demand and uptake, especially costs that involve medicines and supplies. These baseline services were built around the following facilities:

- Ambulances
- Level III health facilities
- Level IV health facilities
- District hospitals (which includes regional hospitals)

Layering SMGL Phase 1 onto the Expectant Mother View

The "expectant mother view" is less apparent in the sections of the model that cost the SMGL program. This is because Phase 1 activities were designed to strengthen and support the Ugandan maternal health care system and not replicate it. Therefore, we built the master process map by linking Phase 1 activities to the Expectant Mother view in the figure below:

Figure 2: Main components of the Saving Mothers Giving Life Program



Most of the activities in Phase 1 were meant to support antenatal visits, delivery and the first 48 hours post-delivery. These are highlighted by the components in the red-dotted rectangle in the above figure.

Findings

The model costs all births in Uganda, broken down by district, and the incremental add-on costs for SMGL, broken down by activities. Based upon the assumptions of the model, we found that the average cost per GoU-only birth (not including set up costs) is $75.94 per birth across the four facilities. This cost should of course be associated with the much higher MMR. In practice, this "average cost" is seldom incurred by most mothers, as it represents the average of normal deliveries and those with complications. More realistic average costs range from $17.63 for normal deliveries at a private Level III, and up to $176 for those with complication.[29]

When comparing these baseline GoU costs with SMGL addon costs, we use the term *improved birth* to refer to the cost of reducing MMR. To reduce MMR under Phase 1's full package, SMGL added $136.12 per birth, yielding a total $212.15 per *improved* birth ($136.21 for SMGL + $75.94 base), which increased the cost/birth by 179% in the first year. These costs ranged from $153.85 for a normal delivery to $312.21 for complications. If we assume that Phase 1 continues to provide the full package into Year 2, start-up costs drop out, and SMGL-assisted births add $85.11per birth to the total, yielding a total cost $162.43per *improved birth* (i.e. almost doubling the cost of a GoU birth).[30] Note that SMGL's cost should be associated with increases to quality and the ability to create and absorb higher levels of demand.[31]

Broken down by activity, the model shows the incremental cost add-on for each activity across the four districts, broken down by year 1 and year 2 costs (the model costs out up to 5 years) [32]:

Activity Break Down of Added Costs Per Improved Birth		
Activity	**Year 1 Cost (w/set up)**	**Year 2 Cost (operational)**
SMGL Management	$18.14	$17.98
Awareness Campaigns	$4.46	$4.17
Village Health Teams (VHTs)	$6.45	$2.80
Boda Transport Vouchers	$3.15	$2.67
Ambulances	$12.96	$1.02

[29] These costs are in part derived by the predicted demand in the four pilot districts. They may shift was the user moves to other districts, and this can be easily seen if the user selects different districts in the model sheet.

[30] The easiest place to find this information in the model is on the CostSum Pilots Tab. We use Year 1 costs columns, which have a 5% increase in deliveries compared to the baseline year, to be able to compare GoU with SMGL Add-on costs, i.e. Year 1 costs for both kinds of birth are derived from the same level of demand.

[31] The assumptions and inputs that define our cost model lead to categories of estimates that are different from those of the Expenditure Study. In the executive summary of the Expenditure Study, the authors suggest that SMGL Phase 1 spent $6.93 spent on capital improvements per birth (p.6).[31] This estimate should not be seen as a contradiction to the estimates produced by the cost model, as they look at different things. The expenditure study estimate looks at the share of capital investments made during Phase 1, approx. $3 million, and assumes a 5-year return on investment with no additional maintenance costs across years 2 through 5. By further assuming a projected 431,276 births across the 4 districts, one can derive the $6.93. Our approach differs in that the cost model also incorporates the other 71% of expenditures that went into Phase 1. Second, our cost model also incorporates year 2 and beyond operating/maintenance costs associated with capital and other expenditures, as well as the depreciation of assets, which is assumed to have an effect on the quality of service delivery. The reason why is because we assumed that the investments made in Phase 1 cannot continue to generate the same level of social returns, i.e. a sustained lower MMR, over 5 years, without additional SMGL or other support. In fact, if we changed the assumptions of the Expenditure Study to allocate the capital expenditure costs only to year 1 (i.e. taking the $6.93*5), the Expenditure Study would predict $34.65/improved birth for infrastructure upgrades, which sits very close to the $30.23/improved birth predicted by the cost model. In other words, the $6.93/delivery should be understood in terms of capital expenditures and the assumptions around the social returns on those expenditures. The cost model transforms a wider set of expenditures into costs, and makes different assumptions about the social returns. As such, we should not conclude that the expenditure and cost model estimates contradict, but should conclude that the former is better viewed as a subset of the latter.

[32] The easiest place to find this information in the model is on the CostSum Pilots Tab. Note, all costs are smoothed out over all facilities in the 4 districts, i.e. 77,184 deliveries (taken from Year 1 on the Scenarios Tab), except for Private Birth Vouchers, which use 13% of the total deliveries as the base and costs out at $78.57/birth. Note also Mentoring and Training do not have separate tabs in the model, but we broke out separately here.

Staff Training	$4.63	$4.40
Staff Mentoring	$2.53	$2.40
Hiring additional personnel	$9.09	$9.07
Private Facility Delivery Vouchers	$78.57	$78.42
Provision of Essential Medicines	$16.13	$17.26
Provision of Essential Equipment	$15.86	$2.46
Upgrades, Renovations and Builds	$30.23	$8.58
System Strengthening	$2.35	$2.20
*Expressed in Cost/Improved Birth, which should be added to GoU Base. Amounts in USD		

It is important to note that these estimates are applicable only to the four pilot districts. They will vary by district according to demand, and the user can select which district, or set of districts (up to every district in Uganda), in the model.[33]

Conclusion

The costing model above answers the Future's Group call "for a more comprehensive understanding of the cost per delivery in each district after all the projected investments have been made and service delivery takes full effect." By reconceptualizing baseline unit cost and SMGL expenditures across a spectrum of maternal health care, we were able to model cost per delivery across varying facilities and districts. This model can also be used to estimate unit costs per birth in other districts based upon varying levels of demand.
Like any model, the results are only as good as the assumptions that lie behind and the inputs that go into it, but the model's structure is made explicit and all inputs are fully customizable according to the latest cost information.

The predictive power of the model is best understood in terms of Pareto's 80/20 principle, i.e. 80 percent of output is produced by 20 percent of input. This means that we aimed to do the first 20 percent of the effort that produces 80 percent of the estimated budget that will be created in Phase 2. The remaining work to be done will involve more detailed cost information on specific line items, staff salaries, etc., but the basic structure and overall estimates produced by the model still provide a useful guide for USAID/Uganda to allocate accordingly.

In the next chapter, we will use the cost model to assign costs to the various pathways identified in the previous chapter to analyze for cost effectiveness. Annexes 3-5 describe how we derived the model, its structure and how the user can use the model for customized cost projections throughout Uganda.

[33] To determine these costs in other districts, the user can go to the district selector tab, select the districts and the changes will be carried through.

Objective 3: Combine the above listed analytical components into a third that compares the relative costs of each combination of activities to determine which are the most cost effective

Findings

Completing Objective 3 involves bringing together the findings associated with the previous two objectives to conduct the cost effectiveness analysis. In doing so, we refer back to the 2x2 metric table discussed in the introduction defined in terms of High/Low levels of effectiveness and High/Low costs. Recall that Quadrant B – the intersection of High Effectiveness and Low Costs – was the preferred quadrant of choice. Reproduced here with the cost and effectiveness findings inserted, we see the following results:

Cost Effectiveness Analysis of Various SMGL Pathways to Reduced MMR
Expressed in $/Improved Birth for Year 1
*Note: Amounts do not include $18.14/improved birth for Program Management Costs

	More Expensive Pathways	Less Expensive Pathways
More Effective Pathways	Pathway 3 - High Demand Pathway: Ambulance & Medicines & Equipment & Staff = $54.04 Pathway 4 - Kibaale/Ambulance Pathway: Ambulance & VHTs & Medicines & Staff = $44.63	Pathway 1 - Level III Pathway: Boda Vouchers & Mentoring & Medicines &VHTs = $28.26
Less Effective Pathways	*Uncommon Pathway* Pathway 2 - Infrastructure Pathway: Infrastructure & Boda Vouchers & Mentoring & Medicines & VHTs = $58.49 *Not very effective* Pathway 5 - Private Facility Pathway: Delivery Vouchers & Boda Vouchers & Medicines & VHTs = $104.30	*Not relevant or effective* Training: $4.63 (not relevant) Hospitals (not effective)

The costs shown above are expressed in $ of each activity path/improved birth for year one and include start up costs. Each pathway should be understood as the add-on costs that are added to the GoU base costs/birth. These costs are also average costs across the various delivery types, i.e. normal and those with complications. They include year one set up costs, but do not include additional program management costs, i.e. $18.14/birth (i.e. add the program management cost to each pathway total).

Based upon the above, it is clear that Phase 2 mostly likely does not need to implement every activity at every facility to successfully reduce MMR by at least 30%. There are two ways to interpret the above CEA results: simple and complex. Under the simple interpretation, the most cost effective pathway forward would be Pathway 1, a VHT/Boda package combined with Mentoring and Medicines, adding $28.26 per birth plus $18.14/birth for program management costs, $46.40/improved birth. Thus, a more cost effective SMGL would reduce add on costs per birth from $136.21 to $46.40 in Year 1, i.e. around one third of the cost of the full blown package, or a reduction of around 66% per improved birth from pilot year costs.

The complex interpretation recognizes how the effectiveness of each program pathway needs to be understood in terms of the context and level of facility within which it travels (as discussed in more detail in the Objective 1 Chapter). For example, the more expensive Pathway 3 may also be considered cost effective if catchment areas witness at least a doubling of facility-based births from the year prior (which occurred around 25% of the time in Phase 1). Alternatively, should Phase 2 decide not to fund Boda vouchers (with or without Every Mother Counts), as in Kibaale, Pathway 4 would be the recommended choice at Level III facilities over Pathway 1.
Moreover, should Phase 2 expand to Level IV facilities, Pathway 2 is a cost effective approach to supporting Level IV facilities that require infrastructure upgrades. Regardless of the pathway chosen, the provision of essential medicines should be part of any cost effective package.

The analysis above suggests that supporting private facilities is not cost effective. The analysis also suggests that training – in its current form – is not needed for Phase 2, and implementers may wish to consider revising the training, or shifting those resources to support more mentoring and staff hiring activities. Finally, the lack of success in Hospitals (also noted in the Objective 1 chapter) may suggest that SMGL's design is less appropriate for addressing the challenges facing those facilities.

Wider Applicability of the CEA Findings and Their Reliability

The first question that should arise when considering these findings is how applicable will they be to other parts of Uganda and elsewhere. On the cost model side, we refer to the 20/80 principle, in which the model's cost estimates should do the first 20 percent of the effort that produces 80 percent of the estimated budget that will be created in Phase 2. Moreover, should the user have more accurate price, input and/or demand information, the model is fully customizable to allow the user to modify the cost estimates accordingly.

On the effectiveness side, the answer to this question requires us to recognize the context-dependent nature of the Pathways identified above. Ostensibly, our findings suggest that some mixture of human resource and material support, combined with some kind of incentive to increase demands for facility-based births, will lead to meaningful reductions to MMR. Yet it may be that in, say, Northern Uganda, facility-based demand is already high and informal networks ensure that essential medicine stockouts never occur. If this were true, then many of the Pathways we found to be relevant in Western Uganda would lose their potency in the North. Conversely, the authors of the Columbia University Final Report do argue that challenges to reducing maternal mortality in the four Phase 1 districts in Zambia are very similar to those in Uganda.[34] As such, conducting a similar CEA exercise in Zambia would tell us a great deal about how far these findings may or may not travel. Yet the point here remains that there are no "silver bullet" or "one-size-fits-all solutions" to the problems facing maternal health in Uganda. If the challenges to reducing the three delays in a given area are similar, and if the contexts within which these challenges occur are also similar, then we do expect that similar Pathways would yield comparable results.

The recognition that some of these assumptions may change or that different conditions may require different solutions should not undermine the reliability of the modeling approach we took to SMGL. This point is better understood in contrast to "expert-based approaches." With expert-based approaches, two different evaluators with expertise in the same area and a comparable level of experience can view the same program and reach very different, even opposing, conclusions. The question of which recommendations to take forward then become a question of how much faith one has in their professional judgments. As such, inter-evaluator reliability is very low. By contrast, we explain, step by step, how we built our models, their assumptions, and what choices we made at critical decision points. Thus, if the same two expert evaluators mentioned above adopted a similar model-based approach, they would not only generate fairly similar results, but they would also be able to compare their choices at critical decision points to understand where and how they differ. As such, the modeling-based approach we've adopted for this exercise does lead to more reliable conclusions.

Conclusions

By way of conclusion, it is useful to ask "how well did this exercise meet its objectives?" We deal with each objective in turn.

How well did this exercise meet Objective 1: Identify which combinations of SMGL Phase 1 activities are necessary and/or sufficient to improve maternal survival in Uganda?

Objective 1 was born out of Columbia University's call to identify the "active ingredients" that define an essential SMGL package. We recognize this as a "managing complexity problem." The managing complexity problem arises when complex programs, like SMGL, need to understand how different combinations of their various activities interact with local contexts in different ways. For most performance management and evaluation teams, the task of identifying trends, patterns, exceptions and conclusions about program

[34] In the first page of the Executive Summary, the authors argue that "the barriers to reducing maternal mortality in Uganda and Zambia include poor quality health facilities, long distances to facilities, a shortage of qualified workers, cultural preferences, and high HIV prevalence. Prior to 2012 in both countries, less than 60% of deliveries occurred in a health facility."

effectiveness across 99 (or even 71) health facilities would be daunting. No doubt SMGL implementers and stakeholders hold a wealth of information about the local contexts, can offer a number of anecdotal examples and possess expertise on what they felt worked or didn't. But the problem they would face involves a lack of rigorous evaluative criteria to analyze this information. Most of the time, this type of exercise becomes an anecdotal reporting of success with limited scrutiny and strong selection bias. Here, there is too much complexity with no method to manage the madness.

Alternatively, statistically-based models - such as head-to-head comparisons[35] - would generate more credible findings due to their reliability, rigor and formality, but would also quickly run into problems involving SMGL's lack of randomized assignment, sample size, ethical considerations associated with control group selection, contamination by other programs, implementation fidelity, Hawthorne effects, etc. Another SMGL-pilot would most likely be required to address these concerns. Moreover, this new SMGL pilot would need to include significantly more health facilities to have a large enough sample size to compare multiple treatment arms (i.e. different activity combinations). In fact, these types of research designs seldom have a large enough sample size (i.e. statistical power) to test multiple treatment arms due to costs. It is also unlikely that a statistically-based model would be able to identify how various combinations of activities may be "path dependent" upon the contexts within which SMGL operates. Instead, the model would simplify results in terms of a one-size fits all solution based upon a regression line, which itself is a single best-fit summary of how variables relate to each other in terms of effect sizes. Here, there is too little complexity to be managed, as it has been replaced with too much simplicity that divorces pathways from their context. Moreover, the single pathway solution gives us nothing to compare it too, so conducting a cost effectiveness analysis of variation *within* the program would be impossible.

As this report has shown, comparative approaches like QCA resolve the complexity management problem by advancing a methodology based upon principles of logical minimization and counter-factual evidence. Rather than view the essential package as based upon a collection of anecdotal success stories, or as the result of expensive control/treatment designs that limit us to simple conclusions of "it worked," the essential package should involve essential *packages*, or pathways, that are characterized by high degrees of "context awareness" (i.e. an understanding of in which set of contexts they are most appropriate). When paired with a cost model, we can also tailor pathways according to higher degrees of cost effectiveness.

Context awareness itself is easier to gain because of the way we conditioned the data to facilitate truth table analysis. To reach the findings, we had to condition the data in ways not usually associated with M&E reporting. Most M&E systems are set up using spreadsheets that put the indicator title at the head of each column and list the beneficiary (e.g. health facility, community, etc.) as individual rows that fall under each column. When reporting, M&E professionals typically look vertically to the top of the column's totals and averages without exploring the variation that lies beneath (as averages, by nature, flatten variation). By contrast, QCA shifts our analytical view to health facilities and their catchment area, i.e. shifted our view from columns to rows. As the findings raised new questions around how SMGL reduced MMR (e.g. how can VHTs be irrelevant when Boda vouchers remain sufficient?), looking at how the various activities combined in different ways at each facility (i.e. looking at rows) helped us find the answer. Moreover, by modeling the implementation in the form of a truth table, we were able to identify contradictions where the same combination of activities led to different outcomes. This in turn compelled us to seek out explanations that invariably involved gaining a better understanding of the local context.

For example, although a district official may know which areas in the district are characterized by more or less adherence to traditional views on community based births, it is seldom clear when/where this knowledge of traditional views will become relevant to programming. But through the approach taken here, this type of knowledge becomes especially relevant when it is called upon to explain why the presence of VHTs in one catchment area still was not enough to convince two women who died of complications at home to give birth in the facility. In other words, QCA-based approaches allow us to couple formal logic with local knowledge to provide rigorous explanations for why context matters. Should a similar approach be embedded within a

[35] The Columbia University PI on p. 5 of the Final Report suggests that "prospective, randomized or quasi-random evaluations, which can be done alongside program implementation, will provide the most credible answers on what constitutes the essential package."

project to support rigorous learning exercise, it would provide a powerful mechanism for conducting mini-evaluations, guiding course-corrections and facilitating the creation of "learning loops" of continuous improvement.

How well did this exercise meet Objective 2: Build a model that can project the approximate costs of implementing SMGL Phase 1 activities elsewhere in Uganda?
We feel confident that cost model we assembled meets Objective 2. Similar exercises of this type often are often forced to humbly conclude that they could've done more with additional data. For us, the conclusion is the opposite, and we arguably had "too much" data that required additional time to digest and prioritize. This success must be attributed to the impressive level of support we received from both implementing partners and representatives at both CDC/Uganda and USAID/Uganda. As such, users are now able to use the cost model to approximate the costs of implementing SMGL's Phase 1 activities elsewhere.

In fact, the user can select which districts, or combination of districts, they would like to use and the model is instantly populated with demand data that combines with input ratios and prices to yield costs. The user can customize costs per district, per activity, disaggregate by facility type, etc. over five years, accounting for both startup and operations cost, broken down by activity. More importantly, even if we prove to be wrong in some of our assessments around, say, staffing ratios, or complication inputs, the model allows those with more accurate information to correct our mistakes and better approximate the true costs.

How well did this exercise meet Objective 3: Combine the above listed analytical components into a third that compares the relative costs of each combination of activities to determine which are the most cost effective?
Finally, fulfilling Objective 3 involves presenting a set of viable alternatives to achieving the same outcome and determining which are the most cost effective. Here, our success is highly dependent upon our success in achieving Objectives 1 and 2, which we have already addressed above. Yet ultimately, the true test of our approach will come if the findings and approach established here are carried forward into Phase 2 and inform the new program design. Should these pathways lead to similar reductions to MMR, and, should their cost effectiveness enable SMGL to reach a wider number of mothers across a wider area, then we can ultimately conclude that this exercise has met its objectives.

Recommendations

Based upon the above, we offer two sets of recommendations regarding program design and program M&E:

Program Design Recommendations

Recommendation 1: Instead of administering all Phase 1 activities to each health facility in Phase 2, SMGL stakeholders may wish to consider following one, or a combination, of the following sets of activities:

- At Level III Health Facilities, esp. hard-to-reach rural ones where new staff turnover is high, offer a VHT/Boda package plus the mentoring of current staff and the provision of essential medicines;
- Should the program be unable to provide Boda vouchers, offer an alternative package of ambulances, essential medicines, VHTs and hire additional staff to support current ones.
- Should Phase 2 expand to Level IV facilities, offer infrastructure upgrades to those facilities that need it, mentor current staff, offer a VHT/Boda package and the provision of essential medicines;
- Should implementers anticipate significantly high increases to facility-based demand (above 50%) in both Level IIIs and IVs, offer an improved and fully-supported ambulance referral system (including the ambulances), hire additional staff, and include the provision of essential medicines and equipment;
- Moreover, consider avoiding Hospitals in Phase 2, or redesigning a different type of intervention specifically tailored to those unique challenges that SMGL is not always able to address;
- Similarly, consider also avoiding private facilities, as they are the least cost effective approach to reducing MMR (due to the high cost of private delivery vouchers) and less effective in general;
- Consider revising the approach to training or shifting human resource supports more to mentoring and/or hiring of critical staff; and
- Crucially, ensure that Phase 2 implementing partners have a heightened understanding of the context in order to identify and adapt which pathway is most cost effective for that context.

Recommendation 2: Consider embedding a similar type of cost effective approach into the ongoing implementation of Phase 2 as a means of periodic review, feedback looping, and continuous improvement.

Recommendation 3: Consider conducting a sister CEA exercise in Zambia, Nigeria and other parts of Uganda to better understand the relative success of each program in general and how far the findings generated here may or may not travel to other countries and areas.

Recommendations based on Lessons Learned from Using Program and Outcome Data for Model Building

These recommendations are largely based upon the discussion in Annexes 2 and 3.

Recommendation 1: In Phase 2, facility names, areas, complications, medicine names, etc. should be standardized according to a consistent protocol. In regards to facility naming protocols, DHMIS names and spacing should be adopted to facilitate easier table merging with other datasets, although program staff should also identify where DMHIS facility names are inaccurate (mostly in regard to misspecification of facility levels). The naming of complications should also follow a standardized format. These standards should be applied to both partner-level activity data as well as survey/assessment datasets, and, if possible, implementing partners should avoid data entry procedures that require staff to hard code these variables into their database.

Recommendation 2: When adopting the facility as a unit of analysis, designate a staff member to consolidate and maintain an overarching truth table. The truth table should identify what activities and resources are delivered to each facility, when and the corresponding amounts (including currency values), for all partners supporting each facility connected to the program (which suggests that a USG staff member may be the most appropriate choice). Doing so requires that implementing partners will streamline relevant program information across various departments (e.g. procurement departments will share the cost of goods delivered, human resource departments will share staff support and turnover data, infrastructure teams will share cost and construction data, etc.). It is also useful to identify thresholds for each activities, i.e. how much of an activity or resource needs to be provided before it is considered "adequate."

Recommendation 3: When supporting an ambulance referral system, partners should develop a logbook that captures both "referred from" and "referred to" fields.

Recommendation 4: At the beginning of the next phase, implementers should identify health facility catchment areas and an agreed upon definition of what constitutes a catchment area. When conducting verbal autopsy surveys, the program should identify an approach for identifying in which catchment area a community based maternal death can be assigned. Enumerators should be trained in these protocols and assign community based deaths accordingly. Ideally, these catchment areas should be standardized and the dataset should be rigorously cleaned to eliminate user entry error.

Recommendation 5: Also in regards to conducting verbal autopsies, implementers should compare their findings with other databases, such as the DHMIS, to increase confidence in the survey's data and flag wide discrepancies that require further exploration.

Recommendation 6: If health facility assessments are conducted (both baseline and endline), assessment designers should identify a means for determining program attribution, e.g. if levels of adequate equipment improve at endline, identify which equipment SMGL provided to attribute the improvement to the program. These assessments should also identify critical staff level of effort at each facility (e.g. part time, full time, concentrated only in maternity ward or timesharing with other wards, etc.) Moreover, questions regarding stockout and equipment levels of adequacy should cover the entire period of performance. Partners should also conduct periodic equipment audits to ensure that the equipment they deliver remains in place and in use. The baseline should be used to construct facility level norms that should guide the level of resources delivered to each facility. These norms should be established before implementation begins.

Annex 1: Scope of Work

Enhanced Cost Effectiveness and Qualitative Comparative Analyses

The USAID/Uganda Monitoring, Evaluation and Learning project (Learning Contract) is the five-year organizational program support vehicle through which USAID/Uganda will pursue realizing its development objectives within the complicated and shifting Ugandan development context. The purpose of the Learning Contract is to provide monitoring, evaluation and learning support services and build capacity of USAID/Uganda to implement USAID's Collaborating, Learning, and Adapting methodology to strengthen the implementation of USAID/Uganda's Country Development Cooperation Strategy and enhance USAID's organizational effectiveness.The Learning Contract has been tasked by USAID/Uganda to undertake an evaluation and a cost- effectiveness analysis of a recently completed maternal health project. The evaluation will utilize the qualitative comparative analysis (QCA) that will be led by an expert in that methodology. We also require short-term technical assistance by a Cost/Financial Analyst expert that can lead the costing and cost effectiveness analysis (CEA) for this activity

Particular areas of interest for the Cost/Financial Analyst will include: Assessing maternal health services costing, cost-effectiveness analysis, efficiency and economic benefits. Assessing long-term opportunities for sustainability financing and potential for integration with other demand-side financing approaches. Working on maternal health costing tools and productions that track and monitor different mechanisms.

Background: The project was designed to accelerate the reduction of maternal and neonatal mortality. Uganda, with a national maternal mortality rate of 310 per 100,000 live births was selected as the second pilot country for this initiative. As a result of the project, Uganda reduced maternal mortality ratio

The project has already undergone an external evaluation conducted by experts. To achieve this success, the evaluation team identified no less than 22 different types of implementation activities including training/mentoring, procuring equipment, supplies and vehicles, construction/renovation, organizational capacity building and partner sourcing, and coordination.

AIM: Overall aim is to determine which aspects of the project's intervention are the most cost effective and scalable in preparation for a national scale up. Illustrative objectives include:

- Provide a detailed costing analysis of expenditures by interventions to more accurately cost out sub activities across its target area.
- Identify which sets of sub activities of the intervention form the necessary and/or sufficient factors associated the project with success.
- Combine the above listed analytical components into a third that compares the relative cost effectiveness of various sets of sub activities to inform the national scale up.

Overall Approach

Traditional CEAs compare the cost of an intervention to its effectiveness as measured in natural health outcomes such as reductions in maternal mortality. While many CEAs compare alternative programs with a common health outcome, this CEA will first and foremost compare various combinations of its sub activities against its own common measures of success. This analysis requires both detailed unit costing information of each sub activity across each facility, as well as rigorous analysis regarding which aspects form the necessary and sufficient factors for success. As such, two inter related methodologies are required to bring together the two sides of costing analysis and program effectiveness.

Costing Analysis Methodology

This is a complex program with many interaction between various stakeholders. This makes it difficult to isolate interventions from each other. Giving these challenges the team is recommending using a business process approach. Essentially the analysis is done from the view of the expectant mother. By following the expectant

mother from the village health team to antenatal visits to delivery and the first 48 hours post-delivery, the team will be able to identify services that the expectant mother receives. Once the different services have been identified, the team can breakdown the services into components and identify the cost associated with those. Interventions could include: VHT, transportation, communications, antenatal and postnatal services, medicines, equipment, etc…

The development of the CEA will require:

- Conduct up to three field visits to randomly selected health facilities (at each level) to understand their "business process" and create a model health facility and processes.
- Meet with district health offices, implementing partners, and other stakeholders to understand the business process.
- Collect and analysis data of implementing partners, government authorities and donors in order to determine costs of each service.
- Calculate the Cost efficiency ratio for interventions using standard CEA approach

Combined CE and QC Analyses

Once QCA has identified various combinations of sets that are associated with positive outcomes, this data can be combined with the costing data to produce unit costs per combination of sets, which can be broken down by such categories as investment, recurrent and community mobilization expenditures across combinations of sets. This analysis should yield an understanding of the optimal combination sub activities and their relative cost effectiveness to other combinations.

Data Sources

The CEA/QCA will utilize previous studies and collected data and collect primary data where necessary

Annex 2: QCA Background and How We Modeled SMGL

This section of the report has been relegated to the annex specifically for those who are interested in learning more how QCA works and the steps we took to reach our conclusions. For the past 25 years, QCA has steadily grown in social science research as a rigorous alternative to statistically-based methods. The method is less known in international development. As QCA moves into international evaluation, adaptations are often required to speak more directly to a policy making audience. Specifically, most QCA exercises are used to explain events and outcomes that have occurred in the past; very few actually attempt to make prescriptions about the future. Moreover, as far as we know, QCA has never been used to inform a large scale maternal health program design as directly as it potentially will under this assignment, creating a potential situation in which the soundness of our recommendations will be judged in terms of the number of lives saved in the next few years. It would not be an exaggeration to say that this is perhaps the highest stress test the QCA approach has faced since its inception.

As such, we retain a natural aversion to recommending a "leaner and meaner" version of SMGL (we call this aversion our conservative bias). We realize that recommending an essential package containing too many activities will drive up costs, and, although it may increase the likelihood that mothers who face life threatening complications will be saved, the added costs of this increased certainty mean that fewer mothers can be saved elsewhere. Yet we also recognize that recommending an essential package containing too few activities is another type of danger. Although it may be able to assist a higher number of mothers across a wider geography, the likelihood that the intervention will lack the "potency" needed to significantly reduce MMR increases. The result is more mothers will die of complications. We confess our bias towards the first scenario – recommending pathways that are potentially "wider" (i.e. comprised of more activities) than need be, and thus our approach favors effectiveness over costs. This bias compels us to make certain adjustments to the QCA findings that are not built into its methodology.

However, we would like to note that should a Phase 2 SMGL incorporate our recommendations, we strongly urge that implementers also incorporate "QCA thinking" into their M&E and Planning systems. Doing so will cut costs, as it will be easier to identify in which contexts narrower pathways can be more effective and in which contexts wider pathways are still required.

Brief Overview of QCA

QCA was first developed by Professor Charles Ragin to help explain social outcomes that lack a sufficient number of cases to use statistical analysis.[1] Since then, QCA has had increasing application in Western policy circles, especially (but not limited to) research inquires based upon medium sized samples (e.g. 15-300). In public health interventions, we see the approach being used to model the relationship between delays to receiving treatment for heart attacks and mortality rates in Brazil[2] to explain reductions in teenage pregnancy rates in England;[3] to understand the conditions affecting early maternal transfer patterns in the US;[4] to explore the relationship between frontline healthcare worker practices and outcomes in Taiwan;[5] to study the factors affecting uptake of new vaccines across Africa;[6] and, more recently, to identify success factors for reducing maternal and child mortality across 144 low- and middle-income countries (LMICs) over 20 years.[7] In many of

[1] Ragin, C. (1987). *The Comparative Method. Moving beyond qualitative and quantitative strategies*, Berkeley/Los Angeles/London: Univ. of California Press.

[2] See Andrade et. al. (2014). "System Dynamics Modeling in the Evaluation of Delays of Care in ST-Segment Elevation Myocardial Infarction Patients within a Tiered Health System." PLoS ONE 9 (7):e103577.

[3] See Blackman, T. (2013). "Exploring Explanations for Local Reductions in Teenage Pregnancy Rates in England: An Approach Using Qualitative Comparative Analysis." Social Policy and Society 12 (1):61-72.

[4] See Britt, D. W. (2006). "A Qualitative Comparative Analysis of the Conditions Affecting Early Maternal Transfer Patterns." Journal of Telemedicine and Telecare 12 (8):392-5.

[5] See Chuang et. al. (2012). "A Configurational Approach to the Relationship between High-Performance Work Practices and Frontline Health Care Worker Outcomes." Health Services Research 47 (4):1460-81.

[6] See Glatman et. al. (2010). "Factors Affecting the Introduction of New Vaccines to Poor Nations: A Comparative Study of the Haemophilus Influenzae Type B and Hepatitis B Vaccines." PLoS ONE 5 (11):1-9.

[7] See Kuruvilla et. al. (2014). "Success factors for reducing maternal and child mortality," Bulletin of the World Health Organization, Volume 92, Number 7, July 2014, 465-544.

these studies, QCA allowed for more granular analysis of how processes of social change work than through using statistically-based modeling.

QCA is used to identify both necessary and sufficient factors that can combine in various ways to produce an outcome. Necessary factors are those that must always, or almost always, be present for the outcome to occur. Sufficient factors are those that, when present, oftentimes lead to the outcome, but other sets of sufficient factors can also lead to the outcome.[8] These notions of necessity and sufficiency are underpinned by set theory, which examines relationships between subsets and supersets. Necessary factors exist as supersets of the outcome, i.e. necessary factors can be present even when the outcome isn't, but the outcome cannot occur unless the necessary factor is present). By contrast, sufficient factors exist as subsets of the outcome, i.e. the outcome can be present even when the sufficient factors are not. This also implies that there can be multiple subsets of sufficient factors. In practice, subset/superset relationships are not always perfect, and QCA has developed parameters of fit to suggest when the relationships are "good enough."

Necessity and sufficiency open the door to the concepts of equifinality and multiple conjunctural causation. Taken together, these terms mean that (1) most often not one factor but a combination of factors will lead to the outcome; (2) different combinations of factors can produce the same outcome (represented as individual solutions in a larger solution set); (3) one factor can have different impacts on the outcome, depending on its combination with other factors and the context; and (4) the absence of a factor may be just as important as its presence.

Extending QCA into development assistance, we deploy it as a comparative approach that integrates qualitative and quantitative methods to account for variations in program effectiveness and explain why impacts occur in some contexts but not in others. For maternal health interventions like SMGL, QCA is able to identify which combinations of activities and other factors are necessary for it to succeed, which ones may not be necessary but are still sufficient for success, and which ones are superfluous relative to the other more essential factors. QCA is therefore an interesting alternative to traditional cost effectiveness analyses, because of its ability to identify the relative effectiveness of various components within a complex program in ways that statistically-based approaches typically cannot.

Crucially, QCA places qualitative "anchors" on the M&E, implementation and other data to make categorical divisions of the presence of a change or non-presence. One way of thinking about how anchors work is to compare it to the Celsius temperature scale. Although it is a continuous scale, the Celsius scale contains two "anchors" – 0 C and 100 C – that represent qualitative changes in the condition of water, i.e. to ice and vapor, respectively.[9] QCA takes this same idea of qualitative anchors and applies it to data around social life. Socially-anchored data abounds in our industry, such as the "extreme poverty" ($1.90/day) threshold the World Bank imposes on daily income. QCA anchors M&E and other data much the same way as the World Bank and Celsius examples do by identifying "thresholds of social change" (we will use the terms threshold and anchor interchangeably). However, many of the M&E and program implementation indicators we use lack clear anchors. When this is the case, we adopt a collaborative approach with technical experts and other stakeholders to identify where to draw substantive boundaries on the data to represent where meaningful change occurs. The advantage is that even if groups disagree on where a certain boundary is drawn, we can easily demonstrate the implications on the results of a wider variety of competing perspectives. What is important is that the investigator clearly identifies what the anchor is and how it was derived.

Once anchored, M&E indicators are recoded as "sets." In addition to inputs/outputs (items within the activity's control), QCA also looks at relevant factors outside of its control, referred to as endogenous and

[8] Certain project activities might be neither sufficient nor necessary but might nevertheless play a role in causing the outcome. These can be called INUS and SUIN factors. INUS is an insufficient but necessary part of a configuration which is itself unnecessary but sufficient for the outcome. Thus, activity A may by itself be neither sufficient nor necessary but may, as part of a combination, have a causal effect. A SUIN-condition is a sufficient but unnecessary part of a configuration of activities that is insufficient but necessary for the outcome. QCA can identify both types of factors.

[9] This example comes from Ragin, C. (2008). "Measurement Versus Calibration: A Set-Theoretic Approach" in *Oxford Handbook of Political Methodology*. Oxford Press.

exogenous sets, respectively.[10] This information is presented in the form of truth tables, which present all the various observed set configurations in relation to the presence/non-presence of the outcome to identify causal configurations of outputs.[11] Truth tables make systematic comparison possible and provide better data visualization. With the truth table in place, QCA then identify how different *combinations of sets* (both endogenous and exogenous) are associated with the presence/non-presence of the outcome across all the facilities affected by the activity, which are called cases. This is done through computer software that applies principles of logic and Boolean minimization to the truth table. The algorithm that QCA uses works according to logical comparisons of presence/non presence conditions to identify minimalized solution sets. A very simple example is below:

A + B + C = Outcome
A + ~b + C = Outcome
Solution: A + C are potential causes of the outcome, but B logically cannot be part of the cause, because we have evidence that the outcome can be achieved without it.

The software returns a set of results that identify which factors (both endogenous and exogenous) are necessary, sufficient, or irrelevant for the outcome to occur. QCA relies upon free software solutions, such as Tosmana and fsQCA, as well as applet-based packages found in STATA and R. For this exercise, we used fsQCA v2.5.

Before moving forward, it should be noted that, in fact, there are several types of QCAs, which differ mostly in the types of variables they use. For this exercise, we use "crisp set" QCA (csQCA), which is based upon dichotomous variables indicating membership/nonmembership (aka present/absent, success/unsuccess, etc.) and represented with 1s and 0s. A more nuanced version of QCA, "fuzzy set" (fsQCA) breaks with the dichotomous variable structure and allows for more nuanced concepts of "partial membership," such as "more a member than a nonmember," and fall within the 0.0 to 1.0 interval. We chose to avoid fsQCA for two reasons. First, it introduces more advanced QCA concepts that may cause considerable confusion for those who are not familiar with QCA. Indeed, the report already takes on the ambitious goal of introducing a new method based upon configurational logic, combining it with a country-wide cost model, and conducting a new type of cost effectiveness analysis. We felt that introducing this additional dimension of analysis may ask too much of the readership. Second, when comparing crisp sets with fuzzy sets, csQCA is more forgiving when it comes to measurement error.[12] fsQCA could be used to recommend combinations of activities that are more precise in terms of prescribed dosage, e.g. not only recommending mentoring (in combination with other activities) but also recommending the degree of mentoring (e.g. how many days). We are not yet comfortable with recommending this level of precision, as it can overprescribe decisions that are better left to the implementers. While we do recognize that embedding fsQCA-type thinking in Phase 2 has potential, we do not feel it is appropriate to introduce such a level of complexity at this stage.

How We Modeled SMGL Phase 1

As has already been mentioned, QCA relies upon logical rules of minimization to identify which sets of "active ingredients" defined various program pathways to reducing MMR. This is done by making counter factual comparisons of examples where a facility reduced MMR with, say, activities A, B and C, but then another facility reduced MMR with only activities A and B (but didn't need C to do it). This "logic of comparison" modeling in practice requires a great deal of variation to be meaningful. In fact, variation is an important requirement for all schools of mathematical modeling; including statistical modeling (e.g. correlations and their corresponding effect sizes can only be uncovered through variation). But unlike statistics, which presents

[10] In essence, QCA reverses the relationship between M&E indicators and interpretation. Most indicators first measure quantities of a given output and then leave it up to M&E and program staff to decide "what the numbers mean." Yet there is nothing inherent in the indicator that says how much of the output is needed to achieve its associated outcome, and thus implementers can conceivably produce many more outputs than what is necessary to achieve acceptable levels of change, as well as not generate enough of the output for a change to occur. By contrast, QCA's anchoring of M&E data not only tells us what the numbers mean, but when they mean something.
[11] QCA has evolved out of its original either/or binary structure (aka crisp sets) into more nuanced and sophisticated approaches involving gradual variations of partial membership/partial non-membership (aka fuzzy sets), combinations of crisp sets and fuzzy sets, and multiple steps involving remote and proximate factors relevant to the success of the outcome.
[12] This point was reiterated by an independent QCA reviewer, who presented the trade-offs of each approach similar to what is presented here.

variation in terms of deviations from some empirically derived measure of central tendency (e.g. the mean), QCA breaks with the laws of central tendency to cluster different groups of variation into various "pathways" of necessity/sufficiency.

This discussion on variation has direct implications for the unit of analysis for this CEA, and explains why we could not use district level aggregates and averages. District averages are only useful if we want to look at variations in SMGL's success across districts, but this would not be very useful as there are only four districts to compare the various combinations of at least 10 different activities. As such, the bulk of our level of effort for QCA was spent breaking down these district level aggregates and rebuilding a model focused on the health facility and its corresponding catchment area as the primary unit of analysis. In doing so, we structured our application of QCA into three overarching analytical steps:

- *Step 1: Model the MMR Outcome and the SMGL Intervention into a Truth Table*
- *Step 2: Conduct Exploratory Analysis to Eliminate Irrelevant Factors, Identify Inconsistencies, and Refine the Model(s)*
- *Step 3: Conduct Confirmatory Analysis to Identify and Consolidate Pathways*

Step 1: Model the MMR Outcome and the SMGL Intervention into a Truth Table

When we refer to "modeling MMR and SMGL," we refer specifically to representing varying levels of MMR and program activities in a spreadsheet. This spreadsheet takes two forms. First, a master data matrix represents all variables or *factors* (e.g. activities, M&E indicators, contextual factors, etc.) we put into the model, and the final column represents the outcome (e.g. changes to MMR). Each row in the matrix represents *case*, e.g. a listing of all the resources that health facility received as well as its associated outcome. Rows of different facilities that have all received the same resources and achieved the same outcome are then grouped into types of cases and presented in the form of a "truth table." Each type of case is presented in the table as a row in which the cells running across the row represent the set of corresponding inputs those facilities received, e.g. how many people were trained at that facility, how many were mentored, if the facility received equipment, etc. The truth table can also identify "contradictions" in which a group of facilities all received the same resources but reached different outcomes (discussed below). In raw form, the content of these cells can accommodate a wide range of variable types – dichotomous, continuous, etc., but in truth table form they are transformed into 0s and 1s, in which a 1 represents membership in a category/or an activity is present, whereas a 0 represents non-membership or the activity is not present.

For SMGL, the primary factors we wanted to model – i.e. include as columns – were SMGL Phase 1 activities and Health Facility-Level MMRs. To move from a raw table to a truth table, we had to apply some kind of criteria – aka a "threshold" – that tells us if that number qualifies for membership in a given category or doesn't qualify. As stated above, membership is represented through 1= membership and 0=non membership denominations. So if a raw number exceeds the threshold, it is designated as a 1; if it falls under the threshold, it is designated as a 0. For example, the threshold for reducing MMR was set at 30%, so a 35% reduction of MMR at a given facility was transformed into a 1 (success); while a 25% reduction was transformed into a 0 (unsuccess).[13] Once we have set the threshold for a factor, it is important that we consistently apply it to the numbers in every cell that fall under a given column. Moreover, each column (variable) requires its own threshold, e.g. a threshold that identifies how many staff should be trained at a health facility before the training activity is considered "present", etc. In what follows below, we identify the steps we took to derive the truth table in terms of the master list of health facilities (i.e. the number of rows), their corresponding changes to MMR (the final column), and the corresponding package of activities each received (the remaining set of columns).

Modeling Facility-Catchment Area MMR Outcomes

Identify SMGL-Facilities to be Modeled

[13] One important side note when constructing truth tables is that a 0 doesn't always translate into the colloquial opposite of a 1, e.g. not being fat does not mean on is thin, not being rich does not mean one is poor, etc. The same applies to the formulation of success – not achieving success does not necessarily imply failure or loss; it just means success was not present. Therefore, we use the term "unsuccess."

Determining how many facilities SMGL assisted during Phase 1 was not an easy task. In terms of the data we collected, we never came across a standalone document that we could use to confidently confirm the total number of facilities SMGL supported. We also found that no one could identify how much support each facility had received, neither within each partner organization nor across them.[14] Thus, a major effort of this exercise was to compile implementation data at the facility level, which we achieved through extensive cooperation with the two main partners. We were unable to identify the additional contributions made by supporting SMGL partners, such as MSH's STRIDES activity or Cardno's SDS activity. We did initially receive a very exciting MS Excel file that consolidated a "planned" 28 activities for 207 facilities across 12 implementing partners/mechanisms, but this was dated July 2012 and was never updated to reflect the actual implementation of Phase 1.[15] Had someone maintained that document, it would have been incredibly useful for all three evaluative efforts of Phase 1. Moving forward into Phase 2, we strongly recommend that someone comprehensively tracks the provision of all SMGL resources to each facility across all contributing partners.

For our purposes, we required health facility names along with their associated levels, parishes, sub counties and districts. In the POMS dataset, we found information for 130 health facilities, 113 Level IIs and IIIs, as well as 17 Level IVs and above, but we were informed that some of these health facilities did not receive SMGL support. In the Health Facility Assessment/Baseline dataset, we found 238 facility entries listed, which covered both a baseline and endline, ideally yielding 119 discrete facilities with two entries each. However, we found that four of the facilities contained three entries, which were government facilities with private wings attached. These contained one baseline and two endline measurements. Based on a comparison of the data, we determined that the baselines closely matched the government endline but did not capture the private wings in the baseline. We therefore dropped the endline private wing rows from the dataset. Additionally, there were four baseline facilities that had no endline, and 15 endline facilities that had no baseline. We dropped these facilities as well, leaving us with a total of 107 facilities remaining. Finally, we went directly to the partners, where we received lists that combined to identify 99 facilities (74 from Baylor and 25 from IDI), 3 of which no longer existed due to closure. These 99 formed the basis of our master list, as implementing partners confirmed their enrollment in the program; the remaining facilities were dropped.

As we compiled the facilities into a master list, we were challenged by the fact that SMGL had not standardized spelling or developed a naming protocol for each facility. For some of the datasets, facility names were hard coded in (as opposed to a standardized drop down list or facility ID number system), and subject to high degrees of data entry error. Some facilities had different names – an official name, local names, nicknames, etc. – and these variations would show up in various datasets or partner lists. Our solution was to cross check the facility naming nomenclature with the one used in DHMIS. However, we also discovered that the DHMIS nomenclature contains errors in terms of facility level (some Level IIIs were mistakenly labeled as Level IIs, etc.). We made note of these errors and standardized our lists according to DHMIS nomenclatures for dataset merging purposes. For Phase 2, we recommend SMGL implementers develop a standardized naming protocol for supported health facilities (or assign a primary ID key) and apply this to all of their associated documentation.

The Outcome Variable - *Identifying the Dataset*

Our next step was to assign MMRs to SMGL facility-catchment areas. This required that we not only identify facility based births and deaths over the baseline to endline period, but that we also capture community based births and deaths over the same time period. In doing so, we searched through various datasets. For this study, we identified three potential datasets that held local area, maternal death information. These were:

- DHMIS Official Statistics: We collected maternal death and delivery data from the DHMIS for the 99 SMGL facilities, as well as for the districts as a whole. We found that for the baseline year, 59 maternal deaths were recorded across the 99 SMGL facilities, compared to 78 deaths during Phase 1. When weighted by the corresponding increase to the delivery rate (46.7%) and standardized to 100,000 births, we find a 9.89% reduction to MMR for Phase 1. When the total number of maternal deaths for all

[14] This was partly due to fragmentation within each organization, as data regarding procurements, trainings, staffing, etc. were distributed across organizational departments, e.g. procurement, H/R, M&E, etc.

[15] The MS Excel spreadsheet's metadata identified the author as "USAID," but unfortunately we never found individual who knew about the document's creation.

four districts was considered, we found a slightly higher 67 and 78 baseline and Phase 1 maternal deaths, respectively, yielding a slightly lower 9.67% reduction to MMR (which makes sense because SMGL did not support every facility in the four districts and thus their contribution lowers the reduction). But these numbers do not capture maternal deaths outside of the facility, and thus understate SMGL's success.

- POMS Maternal Health dataset: This dataset was comprehensive in terms of collecting additional facility level information related to maternal deaths, but this data did not have community-based mortality information, so its corresponding maternal death rate variables (there were several) were not used for the outcome variable. The maternal deaths in this dataset closely, but not entirely, matched those of DHMIS. We did use this data to determine complication ratios.

- Verbal Autopsy Data: We turned to the Verbal Autopsy dataset, which logged 589 deaths across both baseline and Phase 1 years, including a baseline 342 deaths and Phase 1 247 deaths. This figure includes deaths in non-SMGL supported facilities as well as community based deaths. We determined to use this dataset to construct our outcome indicator (see below) because 1. it captured community maternal mortalities, and 2. it had already been used to inform previous SMGL reports.

Assigning Deaths to Facilities and Defining the Catchment Area

After dividing the maternal deaths into base and Phase 1 time periods, we sought to identify which fatalities could be assigned to SMGL-supported facilities. The dataset indicates that 327 maternal deaths took place in some kind of health facility, implying that 262 of the entries were community-based deaths.[16] In order to apply QCA and conduct counter-factual comparisons, we needed to disaggregate the 30% reduction figure to facility-catchment level MMRs and assign community-based deaths to a facility's corresponding catchment area.

Defining the catchment area essentially revolves around the question of which facility the mother *should have visited* to seek treatment and support.[17] We identified two competing criteria regarding how to identify where the mother should have sought care. In the first approach, one could define the catchment area in terms of the facility closest to where the mother lived that was able to perform BEmONC services. In many cases, this assumes that the mother *should have visited* a Level III facility first, unless a Level IV or Hospital happened to be closer. Should complications arise, the mother would then be transferred to a higher level facility. Alternatively, one could assume that all community-based maternal deaths involved complications that should've been addressed immediately at CEmONC services. We would then define the catchment area in terms of the closest Level IV facility or Hospital, and assume that the mother should have bypassed the closet Level III altogether.[18] There are strong arguments for either decision:

- BEmONC-defined Catchment Areas: One of the main goals of SMGL was to encourage Ugandan mothers to shift away from non-facility based deliveries towards facility based ones, while at the same time reduce the burden on the already overloaded Level IVs and Hospitals by shifting normal and less complicated cases to BEmONC facilities. VHTs and communication outreach activities were designed to facilitate this shift, and SMGL also upgraded many facilities to BEmONC in order to improve the health system's credibility. Indeed, implementers chose to organize VHTs around lower facility level-

[16] The verbal autopsy survey asked respondents where the expired mother died and, if in a facility, the name of the facility. These variables are found in the INF309 series of the Verbal Autopsy dataset.

[17] In situations where the mother died on the way to a facility, the question is even more complicated – should she have decided to leave sooner? Or did she decide to attend a facility that was too far and may have been saved had she decided to travel to a closer facility? In cases where the respondent indicated to which facility she was travelling, we assigned the death to that facility, but this only occurred in a few cases.

[18] We also identified a third, hybrid set of assumptions that assumes the mother herself, perhaps as a result of referrals for more specialized she received during her antenatal care visits, was best able to define what level of EmONC she should receive prior to her journey to the facility, and therefore the catchment area could be defined in terms of this decision. One way of inferring what her decision was would be to examine each autopsy report on a case by case basis, and define her catchment area (BEmONC- or CEmONC-based) in terms of type of complication. We didn't feel comfortable with this approach as it 1. was beyond the skillset of members of our team; and 2. was the approach most open to individual-based interpretation, as we were informed that two different medical officers could reach different conclusions over where to send the same patient. We therefore didn't pursue this option, and focused on the other two.

defined catchment areas. Also, SMGL's approach to strengthening the ambulance referral system involved prioritizing the transfer of patients from BEmONC to CEmONC facilities over home-based transfers, thus implying that mothers should first attend BEmONC facilities before being transferred. As such, assigning community-based deaths to the BEmONC catchment area will help us better understand to what extent these activities were working.

- CEmONC-defined Catchment Areas: Alternatively, one could argue that VHTs also sought to encourage mothers to visit facilities for antenatal care, where potential complications could be identified early on and mothers would've known that they should proceed directly to facilities that provide CEmONC services at the onset of early labor. Using this definition, we would be better able to evaluate how successful SMGL was at establishing an effective early warning system.

Given that the definition of the catchment area was based upon hypothetical decisions informed by two different sets of assumptions, we chose the one that was most conducive to program learning in line with the objectives of our scope of work, i.e. to identify cost-effective pathways for the Phase 2 rollout. At the beginning of the exercise, we were informed that SMGL was unique in its ability to increase demand for BEmONC services at the Level III facilities, and that this approach was more likely to be taken forward in Phase 2 than the more resource intensive support required at the level of the Hospitals. Additionally, adopting a BEmONC-based catchment area (save in cases where a Level IV or Hospital was closer) would increase the number of facilities with changes to MMR that we could model. This in turn would lead to more diversity across the truth table and strengthen the refinement of the various pathways to reducing MMR. Conversely, there was less to be learned through defining the catchment area solely in terms of CEmONC services, as it would've only increased MMR at the Hospitals (they would've been assigned more deaths, increasing the numerator), highlighting the relative lack of success at those levels even further, and leaving many questions about SMGL's effectiveness at the lower levels unanswered.

We therefore defined our main unit of analysis as the health facility and its corresponding catchment area. In most cases, the catchment area was defined in terms of the closet facility that could provide BEmONC services, except in cases where a CEmONC facility was closer. In doing so, we followed the catchment area system used by the implementing partners to assign VHTs to communities and the facilities they recommended mothers use at the onset of labor.

Identifying SMGL-supported Facilities with Changes to MMR

Once this decision had been made, we undertook the labor intensive process of assigning the 262 community based deaths to facilities and their corresponding catchment areas. This involved the previously mentioned challenge of standardizing facility names, identifying the nicknames entered by enumerators, etc. However, along the way, we made two additional discoveries regarding the alleged 327 facility based deaths: First, eighteen of the so-called facilities in fact where not officially-recognized facilities, but were either listed as undecipherable vendors, e.g. "Teddy's drug shop," "private clinic in Kyenjojo," etc., or miscoded as "N/A", "on the way," etc. These should not have been counted as facility based deaths and used to determine the change to facility-based MMR. We reassigned them to corresponding catchment areas based upon the village locations identified. Second, ten of the deaths occurred in Entebbe and/or Mulago (in Kampala) because the mother had moved, had visited prior to the onset of labor (up to a month prior), or had visited these areas after returning home from delivery (up to two weeks after birth). We assumed that these cases should not have been included at all in the overall change to MMR across the four districts, and did not assign any of these cases to a local catchment area. Once this had finished, we identified 527 deaths that could be assigned to SMGL-supported facilities and their corresponding catchment areas, while 62 deaths were removed.

To ensure our numbers were accurate, we cross checked the deaths identified in the verbal autopsy data with those in the POMS and DHMIS datasets. In the POMS data, we found three additional health facilities with maternal deaths that were not detected by the verbal autopsy survey (Nyantaboma HC II, Kasusu HC III, Kagote HC III). We added these deaths to our table. We also found an additional death in the DHMIS data that was not found in either the POMS or Verbal Autopsy datasets (Kicheche HC III).

For the denominator, we faced the challenge of a lack of reliable information on community based births. DHMIS only collects reliable information on facility-based births, and the verbal autopsy survey excluded community based live births (they only focused on deaths). Were we to only use facility-based births in denominator, it would have over-stated the MMR. We discussed this issue with CDC representatives and agreed to compensate for the lack of reliable community-based births using Demographic Health Survey-based age specific fertility rates.[19] This data was used to make estimates for the total number of live births across all four districts and formed the denominator used to calculate the 30% reduction in MMR. We then aggregated the total number of live births found in the DHMIS database across all four districts and found that the DHS estimates exceeded DHMIS records by around 15%. We attributed the difference between the two numbers to community-based births. We then applied this adjustment to the denominator of facility-based births to define the total number of births (both facility and community) in that facility's catchment area.

We standardized these rates as ratios by weighting to the 100,000 deliveries used in MMR. Because we needed to identify *changes* to maternal mortality rates over time, we divided the deaths according to baseline and Phase 1 time periods. Out of the 99 SMGL-supported facilities and their corresponding catchment areas, we found 71 with changes to maternal mortality numbers from the baseline to Phase 1 year (the rest had no maternal deaths in either years, primarily because most of them are Level IIs that are not designed to offer delivery services).[20]

The final step in defining success was to assign a threshold of success to determine how the facilities would be presented in the truth table, i.e. was success "present" (1) or "not present" (0). Considering how SMGL stakeholders have already embraced the 30% reduction as a success for Phase 1, we adopted this figure as the threshold defining membership in the success category. Across these 71 facilities, we found:

- 56 facilities reduced MMR by more than 30%, and there was an overall reduction in MMR of 48%
- 15 facilities did not pass the 30% reduction threshold, as 10 facilities witnessed increases in MMR (with an overall average increase to MMR by 56%) 5 additional facilities reduced MMR, but did not succeed in reducing it by at least 30%

The 71 health facilities analyzed should not be understood as a sample with a corresponding probabilistic degree of general representativeness. Rather, these 71 facilities represent the segment of the SMGL-supported health facilities that had varying levels of maternal mortality that could be modeled. The implications are that the above listed descriptive statistics should not be understood as standing in conflict with other findings in other SMGL reports, but should be understand as different results that correspond to adopting different levels of analysis. It is not surprising that our approach found a higher overall reduction of MMR in the 71 facilities than across the four districts as reported in the Monitoring Mothers report, because the latter figure also included unsupported facilities. From the "SMGL-only" perspective, Phase 1 was actually more successful than was previously reported when we shift the unit of analysis. Moreover, shifting the unit of analysis to the health facility also allowed us to bring problem facilities into clearer focus; something which is more difficult to do when the unit of analysis is on district totals.

It should also be noted that there was some concern over the high degree of sensitivity associated with using facility-catchment area MMRs. For example, smaller catchment areas that have a lower number of births, small changes to the maternal death rate (e.g. 1-2 deaths) become amplified when transformed to the 100,000 birth ratio. While this concern is valid, it is to some extent mitigated by the fact that we transform these MMR results into sets for QCA, and thus sensitivity only becomes an issue when an MMR crosses above or below the 30% threshold. For example, if a facility catchment jumped from a 0% to a 29% decrease in MMR, it would

[19] See the Maternal and Perinatal Outcomes in Health Facilities report, found online at http://savingmothersgivinglife.org/docs/Maternal_Perinatal_Outcomes.pdf

[20] Facilities that witnessed deaths in during both baseline and Phase 1 periods, as well as during the baseline but not in Phase 1, were easily assigned percent changes to MMR. However, we also found 10 facilities that had no maternal deaths in the baseline year but did under SMGL Phase 1, which lead to a problem of calculating a rate of change with a zero in the denominator. For all 10 facilities, we searched DHMIS for additional maternal deaths that may have appeared prior to the baseline year (back to 2009), but found nothing. We were informed by program staff that most likely these facilities had maternal deaths, but they were not reported as the earlier HMIS system was not effective. As such, we assigned the overall MMR baseline value of 492 deaths/100,000 deliveries as the baseline for each of these 10 facilities as a substitute to derive their changes to MMR.

not change how that catchment area was scored. We tested this by artificially adding/subtracting deaths to each facility catchment area, and found that only four were affected. Thus, although there is still a degree of sensitivity at the 100,000 births level, we don't find this sensitivity to be significantly threatening to the overall results.

Modeling Phase 1's Activities

Modeling the SMGL intervention was a relatively straight forward task. Drawing upon the Columbia University evaluations, we reviewed all 24 activities offered by the program (see the activities chart in the QCA chapter). We then consolidated many activities – such as the various trainings – into master activities, to identify 10 categories of Phase 1 activities that could be modeled at the facility level:

- Presence of Village Health Teams (coded as vht)
- Facility Upgrades and Renovations (coded as infra for infrastructure)
- Vouchers for Boda Rides to Deliver at Facilities (coded as vbod)
- Vouchers for Deliveries at Private Facilities (coded as vbir)
- Referrals to Facilities brought by Ambulances (coded as ambu)
- Critical Staff Hired in Facilities (coded as staff)
- Facility Staff Mentored (coded as mento)
- Facility Staff Trained (coded as train)
- The Provision of Essential Medicines (coded as supp)
- The Provision of Essential Equipment (coded as equip)

We obtained equivalent data on the implementation of these activities from Baylor and IDI. We could not model system-strengthening activities, such as district level capacity building efforts, as well as some awareness raising activities, such as radio programs, because they could not be assigned as facility/catchment-level attributes. In what follows, we briefly describe the activities we modeled, the datas source we used, and the thresholds we applied. We worked closely with both implementers to identify reasonable thresholds, cross-checked these with both USAID and CDC representatives, and ran several iterations of the QCA model with different thresholds to identify which factors were more sensitive to threshold setting. Where QCA found significantly different solutions, we noted the sensitivity of these thresholds.

Staffing (threshold: at least one full time staff member)

Staffing data was collected from both partners. To model additional staff hired by SMGL, we needed a method to assign relative weights to various types of staff categories, so that the contributions of, say, hiring a medical officer can be compared with those of hiring a midwife. To do so, we modeled partner staffing contributions in terms of their percent contribution to the total essential personal workforce, in which "essential" refers to Medical Officers, Nurses, Midwives, Clinical Officers, Anesthetic Officers and Lab Technicians. We first needed a baseline of the average essential staffing levels across the various health facility types to identify the "normal" staffing levels. In doing so, we surveyed 40 health facilities and asked respondents to quantify their contributions in terms of labor hours because many essential personal do not work full time at any given facility (implementing partners did not track labor hours and could not identify who was part time/full time). Based upon the survey, we identified various labor hour norms, per category and per facility type. We then weighted the SMGL partner contributions by these norms and derived a composite score. This is show below:

Staff Level of Effort Norms Based upon SMGL CEA Survey of 40 HFs, expressed in # of full time staff per facility							
	Doctor/ Medical Officer	Nurses	Midwife	Clinical Officer	Anesthetic Officer	Lab Tech	Total Staff Score
Level 5	2.8 (18.4%)	3.75 (24.6%)	5.8 (38.0%)	1.1 (7.2%)	0.9 (5.9%)	0.9 (5.9%)	15.25 (100.0%)
Level 4	2.28 (19.7%)	0.67 (5.8%)	6.4 (55.4%)	0.28 (2.4%)	0.33 (2.9%)	1.6 (13.8%)	11.56 (100.0%)

Level 3	0.04 (0.7%)	1.22 (21.9%)	3.08 (55.2%)	0.64 (11.5%)	0 (0.0%)	0.6 (10.8%)	5.58 (100.0%)

Each facility level received its own composite score, and is expressed as the total average percentage staff contribution for that facility. For example, hiring a fulltime midwife to support maternity at a Level IV facility would result in a 8.7% overall staff increase.

During Phase 1, we found 19 of the 71 health facilities did not receive any staff contributions, 22 had very small contributions of only a part staff member, and the remaining 30 had full time contributions of one or more staff member. As such, we set our minimum threshold for the presence of staffing support at the level of at least 1 fulltime staff member (1) of any category, and assigned a "not present" (0) to facilities that received no or only part time staffing support. The threshold is relatively stable and the results don't change significantly until we raise the threshold to four fulltime staff.

Mentorships (Threshold: Any mentoring present)
We modeled the mentorship activity in terms of the number of mentoring sessions conducted per facility. Partnership data allowed us to identify the facility of the mentored. The mentoring activity was spread out across seven sessions in total, in which one employee could benefit from multiple sessions:

Distribution of Employees receiving multiple mentoring sessions		
#	Frequency	%
1 visit	619	82.55%
2 visits	74	11.95%
3 visits	26	4.20%
4 visits	6	0.97%
5 visits	2	0.32%
Taken from IP data		

In fact, we find that the norm was for most employees to attend only one session, with very few instances of an employee attending 3 or more. We also found that while 11 facilities hosted 5 or more mentorship sessions (including Level IIIs), the norm centered more around 3 mentoring sessions, with almost half (35) of the 71 health facilities with no mentored employees. Given this complexity, we modeled two different scenarios to determine if mentoring was "present" or "not present" in the truth table: Relaxed, i.e. if any mentoring session was conducted at a health facility, it received a 1; as opposed to Moderate, i.e. a health facility received the "norm" of at least 3 mentoring visits. We found that running either scenario in the model didn't really matter for the overall findings presented in the previous chapter, save in the case of Pathway 4, which involved infrastructure activities. There mentoring dropped out as irrelevant when we applied the moderate scenario (there was counter-factual evidence that relaxed mentoring wasn't needed), but emerged as relevant when we applied the relaxed mentoring scenario of any mentoring as present. Given that this activity was threshold sensitive at these levels, we included it on Pathway 4 (see below) as a conservative measure to ensure the width of this pathway was accurately defined.

Trainings (Threshold: at least 2 staff trained at a facility)
We modeled the presence/non presence of training through an approach similar to mentoring, in which we counted the number of trainings administered to various essential staff at the health facilities. The training data we received from the partners identified 278 trainees across three subject areas: ALARM/newborn training (182), PMTCT/EID training (45), and comprehensive training of trainers training (51). We assigned equal weight to each one of these trainings and represented their level of intensity in terms of the number of employees trained at each health facility. The total number of employees trained per facility ranged from none to 19 (at Kagadi Hospital), with an average of almost four employees trained per facility. We again set the threshold on various scenarios (relaxed = 2 or more trained, moderate = 3 or more, and strict = 4 or more). However, we found that none of these thresholds mattered when we ran the model.

Equipment (Threshold: "adequate equipment" in at least 2 of 6 functional areas)
The provision of equipment turned out to be a very difficult activity to model. We did receive an extensive amount of procurement data from both partners that identified what was procured.[21] However, we were unable to establish a means of standardizing the relative weight of this equipment as we did not receive the associated cost information for these procurements.[22] Baylor's data did contain some cost information on equipment, but this was lumped together with a variety of procurements made by different projects and kept in a massive folder of data that was difficult to extract. IDI, on the other hand, informed us that they did not have their procurement cost information in electronic format, which suggests that they were unaware of how much they were spending on equipment. Apparently, IDI allocated purchases for equipment according to bulk deliveries, but did not track unit costs (which vary per delivery) within those deliveries. As such, there was no way of knowing how much IDI spent on equipment per facility, or how much they added to what was currently there, so we could not use this approach to model equipment.

We instead turned to the Health Facility Assessment, which was administered both at the beginning and end of Phase 1. The Health Facility Assessment asked respondents if there were adequate levels of equipment present in the labor ward for emergency obstetric care. Specifically, we pulled the questions that asked if equipment to perform vacuum extraction, manual vacuum aspiration, maternal and new born resuscitation were consistently available and functioning, as well as the presence of full infusion sets and fully stocked emergency trollies. In this instance, we calculated the results in terms of changes in adequacy from the baseline year to the Phase 1 year, as opposed to simply aggregating the number of facilities with adequate equipment in the Phase 1 wave. The rationale behind this is that the baseline levels of equipment were associated with a much higher MMR. Here, we were trying to understand how the intervention reduced MMR, which implies that we should capture levels of improvement vs. overall levels of adequacy. We then constructed a composite score based upon these questions to give an overall level of adequate emergency equipment present in the labor ward for emergency obstetric care. Fourteen of the 71 facilities logged no improvements to equipment adequacy, 10 logged only one improvement, and only three managed to log improvements in 5 of the 6 areas. Finally, we assigned the threshold that at least two of the six functional areas were required to be adequate for the provision of emergency equipment to be considered present (1) in the truth table.

Essential Medicines (Threshold: A lack of stockouts from at least three of the five categories of essential medicines)
Similar to the provision of equipment, we originally sought to model the provision of essential supplies in terms of their ongoing monetary contribution. However, as was the case of equipment, there was no baseline data for us to use to understand SMGL's contribution in relative terms. We sought to monetize the value of essential medicines provided by SMGL. However, we were unable to establish a means of standardizing the relative weight of these supplies as we did not receive the associated cost information for these procurements in the case of IDI, or confirmation on where these medicines were used. As in the case of equipment, Baylor's data did contain some cost information, but this was lumped together with literally thousands of supply purchase items made by different projects and kept in a massive folder of data that was difficult to extract. IDI, on the other hand, did not have any supply data in electronic format. We were informed by the Kampala office that this information was kept in the field office and tracked through hardcopy requisition forms. However, when we arrived in Kibaale, these forms could not be located and we were informed they were kept in Kampala. IDI did finally locate the forms during the last few days of our visit and we set our survey team to convert the supply information into electronic format. However, the requisition forms they provided were very difficult to decipher, many lacking dates, etc. It was also unclear to which facility department (e.g. maternal, HIV, antenatal, etc.) these medicines were delivered.

We used data from the Health Facility Assessment that asked respondents if there were stockouts for medicines required in the labor ward for emergency obstetric care. Specifically, we pulled the questions around stockouts in antibiotics (including Ampicillin, Gentamicin, Metronidazole, and Procaine benzyl penicillin G), Magnesium

[21] According to the equipment delivery data we received from the implementing partners, the top seven types of equipment SMGL provided to health facilities (in order of most common to less common) were delivery sets, dispensing trays, drip stands, autoclaves, caesarian sets, suction pumps and oxygen concentrators.
[22] Note, our requirements here were different from those of the costing. For the costing, we only required enough procurement data to derive estimated average costs for the equipment. By contrast, for the QCA we needed to understand actual expenditures on equipment by facility.

Sulfate, Oxytocin, Atropine, and Ketamine. Although there were other medicines that could've made this list, the Health Facility Assessment only asked about stockout levels for these five across the entire Phase 1. We then constructed a composite score based upon these questions to give an overall level of consistent emergency medicine presence. Only one health facility had no stockouts of any essential medicines during the year (Bukuku HC IV), only two witnessed stockouts in only one category, even as 15 out of the 72 facilities had stockouts in all essential medicines identified. The rest were concentrated around stockouts in between two to four essential medicines. We first designated that a lack of stockouts for three out of the five essential medicines needed to be present to be considered adequate in truth table, but found that our model results didn't change when we moved the threshold between two and four out of five markers. We kept the three out of five threshold marker in the model following the guidance of the Baylor Medical Officer.

Infrastructure (Threshold: Any infrastructure upgrade/renovation was present, but didn't include placenta pits)
Modeling infrastructure was a relatively straight forward process involving the identification of all health facilities that received physical upgrades, renovations, constructions, etc. This data was captured in both partner records as well as the Futures Group study. We logged any form of upgrade as "present," save the construction of placenta pits (as per partner advice), while the rest remained "not present."

Boda Vouchers (Threshold: at least 10% of total deliveries at the facility involved mothers using Boda Vouchers)
We modeled Boda Vouchers in terms of the number of vouchers received at each facility weighted by that facility's corresponding number of deliveries. We received Boda voucher data from Baylor and compiled the total number of vouchers received at each facility. We then divided the number of deliveries at the facility into the total number Boda voucher assisted births. Twenty seven facilities did not log any deliveries assisted with Boda vouchers (i.e. they were from IDI), 10 facilities logged less than 10%, but the majority of facilities logged Boda voucher assisted deliveries ranging from 20% to 100%. For Boda vouchers to be considered adequate, at least 10% of the total deliveries needed to involve Boda vouchers. Note: many women who give birth at Level IV facilities and above are from urban areas and would not require a Boda voucher in most cases.

VHTs (Threshold: at least 10% of total deliveries in the facility involved mothers who were supported by VHTs)
Similar to Boda Vouchers, we modeled VHTs in terms of the number of VHTs registered at each facility weighted by that facility's corresponding number of deliveries. Both partners assigned VHTs to a discrete facility catchment area (i.e. no VHT was designated to more than one facility), and we aggregated the total number of VHTs across each facility. We then divided the number of deliveries at the facility into the level of VHT presence to derive a VHT presence to delivery rate. Sixteen SMGL-assisted facilities did not have any VHTs assigned to them, while four facilities registered more VHTs than births (e.g. in Kagote HC III, there were 36 assigned VHTs, and only 4 births). We set the same presence threshold of at least 10% VHT presence to deliveries in our truth table, but this time it was due to the recommended practice of one VHT worker to ten expecting mothers.)

Delivery Vouchers (Threshold: the presence of any delivery voucher used at that facility)
We modeled delivery Vouchers the exact same ways as we did for Boda vouchers in terms of the number of vouchers received at each facility weighted by that facility's corresponding number of deliveries. Nine private facilities logged delivery vouchers, which covered from a little more than 50% to almost 98% of deliveries at these facilities. Given the high level of presence, we logged any private health facility that claimed delivery vouchers as present (1), while the remaining were not present (0).

Ambulances (Threshold: any high risk mothers who arrive at a facility by ambulance counts as present)
We again took the weighting by birthrate approach to modeling the ambulance data, i.e. by weighting the number of mothers delivered by ambulance as a percentage of the total number of deliveries. However, we were only able to do this for the "referred to" data, i.e. which lists the facilities receiving mothers (they inherent the risk), and not the "referred from" data, which lists the sending facilities (and they shift the risk). We were unable to model referred from data because IDI did not capture this information until the end of Phase 1 and thus our dataset was incomplete. We found that 36 facilities received mothers with complications via ambulances, which ranged from .1% to 28% (at Kabarole Hospital) of all deliveries at the facility. Given that the arrival of a high risk mother with complications can have a direct impact on maternal mortality rates, we

modeled in the presence of ambulance referrals at any facility that received a patient, e.g. .1% = present (1), and only facilities with zero receiving referrals were logged as not present (0).

Finally, we also added facility type (government, or private), facility level (i.e. Levels III – Hospitals), implementing partner (Baylor or IDI), district and delivery rates into the model. This gave us a total of 15 factors and the MMR outcome that defined the truth table across 71 facilities.

With these thresholds in place, we were able to build a master data matrix that identified all 71 facilities, the resources they received/didn't receive, their corresponding outcomes, and additional descriptive factors. This master table formed the basis of the models we subsequently tested and refined, described below.

Step 2: Conduct Exploratory Analysis to Eliminate Irrelevant Factors, Expand Coverage, and Refine the Model(s)

After the data matrix and corresponding truth table was built, we conducted several iterations of exploratory model running to find the best fit for the data and seek the most parsimonious set of solutions (QCA does not typically return just one set of solutions, but many). One of the main goals was to identify which factors were relevant and which were not. Many QCA enthusiasts argue that the method is ideally suited for small to medium sample sizes. While this is true, the QCA software (we used fsQCA) is still limited in the number of factors it can efficiently analyze, for three reasons. First, fsQCA usually faces processing constraints after 11 factors; more factors can lead to long processing delays or even crash the software. Second, more factors substantially increase the size of the truth table, which is computed by 2^n, where n=the number of factors, e.g. 11 factors leads to a truth table size of 2,048 rows. Yet in practice, we observed only 71 facilities, some of which are further combined into "types of cases" in the truth table (up to 58, depending on which factors are included). This leaves a substantial number of unobserved combinations of resources, aka "limited diversity," which can limit our ability to draw conclusions. Third, increasing the number of factors increases our vulnerability of what can be considered the equivalent of a Type I error: accepting that there is a relationship among one or more factors to the outcome when in fact they are instead artifacts of the way the data is structured. One way of managing these issues is to reduce the number of potential factors so that the overall truth table size is smaller, but this should only be done if there is a substantive reason to do so.

To assist, we first ran an exploratory model to identify how well and to what extent the various solutions offered explain the patterns observed in the data. In doing so, we explored how to improve two main parameters of fit in the model: Consistency and Coverage. [23] Consistency assesses the strength of the relationship of the solution offered to the outcome. With crisp sets, consistency can be seen as the proportion of cases with a given causal combination that are also in the outcome set (it is more complicated with fsQCA). Identical cases that can be associated with both the outcome as well as the absence of the outcome can be referred to as "logical contradictions," lowering the consistency score of that individual solution as well as the overall consistency of the model. Contradictions in crisp set analysis oftentimes require further investigation. Best practice suggests that solutions and solution sets with consistency scores lower than .8 should not be accepted, as they indicate that the solution does not regularly lead to the outcome.

Coverage is a measure of the empirical relevance of both each individual solution as well as the overall solution set. Although a particular solution may be perfectly consistent with the outcome it explains, the number of facilities it explains may in fact be very low. Coverage scores are divided into raw and unique coverage, in which the former can consist of facilities that could be assigned to more than one solution, while unique coverage refers to the coverage of facilities that are only covered by that solution. There may even be solutions that contain no unique cases but could be entirely explained by other solutions. Finally, solution coverage refers to how much of the outcome is covered by the entire solution set. When conducting an exploratory analysis, we dropped terms that QCA eliminated as irrelevant, as well as dropped terms that had no impact on the coverage and consistency scores.

[23] Consistency scores can be likened conceptually (but not mathematically) to significance values in inferential statistics; while coverage scores may be likened to r-squareds (for overall coverage) and correlation coefficients (for individual pathway coverage). However, see Schneider, C. and Grofman, B. (2006). "It Might Look Like a Regression ... but It's Not! An Intuitive Approach to the Presentation of QCA and Fs/QCA Results," Compasss Working Paper, WP2006-39/

We ran several iterations of various activity combinations, iteratively removing and replacing each factor to understand if their impact on the consistency and coverage scores. We also repeated this exercise with fluctuating threshold markers to identify which thresholds were the most vulnerable to mild fluctuations (discussed above). Throughout many of the iterations, a persistent pattern of contradictions occurred. Although contradictions, in and of themselves, are not necessarily problematic (as long as the overall consistency remains above .8), but they are nonetheless still worth exploring. We initially found seven contradictions in the data. Upon closer inspection, the first set of two contradictions involved pathways that had both private facilities and government ones. A second set of two contradictions involved pathways that included Hospitals and Level IVs. The three remaining contradictions were:

- For reasons we cannot explain, Nyamabuga HC III was in contradiction with other facilities that received the same intervention (Kaswa HC III and Nyankwanzi HC III, both of which were successful). Although Nyamabuga received relatively the same level of resources from SMGL, they had an increase in maternal mortality from zero deaths during the baseline to three deaths in Phase 1. The increase in deliveries was relatively low at 27.8%. Upon closer inspection of those deaths in the verbal autopsy data, we found that all occurred outside of the facility and respondents referred to witchcraft as the main cause. This may imply that the VHTs assigned to this area were not able to overcome strongly held traditional beliefs, were not able to reach these women in general, or a host of other factors. We couldn't resolve this contradiction by introducing another variable or lumping it with a sub model, and we didn't have a theoretical justification for removing it, so we left it in the model to highlight that unique contextual factors that run counter to wider trends are always a possibility, but QCA nevertheless provides the clues to identify these exceptions and why they matter.
- We pulled Hospitals from the model. The types of complications experienced at Hospitals tend to be qualitatively different and more difficult than at the lower level facilities, as witnessed by the fact that MMR actually increased in these facilities under SMGL. Indeed, it may be that MMR increased *because* of SMGL, as the overall improvements to the regional referral system sent more problematic and terminal cases to Hospitals when they might have previously occurred at lower level facilities. Moreover, the level and type of resources needed to enhance hospital quality to reduce MMR may be significantly larger than those offered by SMGL. For these reasons, we deemed it appropriate to pull hospitals from the model, on the grounds that their processes are too different to compare with the lower level facilities.
- After several iterations, we also pulled private facilities from the model, and ran a separate private facility model, which defined Pathway 5. This decision also meant that we would exclude private delivery vouchers from the revised, government only model, as they were only relevant in the private facility context. Our main justification for this is that SMGL operated differently in private facilities than in government ones. In the latter, SMGL oftentimes directly provided the resources to the facilities, while in the former, implementers assumed that private facility management would reinvest the profits they gained from private delivery vouchers back into the facility in such things as additional midwives, more medicines, etc. Our interviews with private facility staff revealed that some private facilities chose to ask staff to stay longer and work harder to meet the increased demand, but did not provide additional resources to support it. As such, direct cash support appears to operate differently than in-kind support, and we feel justified in breaking out private facilities as a separate model.
- Finally, as suggested in the main body of this report, QCA identified training as an irrelevant across the various iterations of combinations we ran.[24] We pulled this factor from the model.

Step 3: Conduct Confirmatory Analysis to Identify and Consolidate Pathways

Before going forward, it is useful to review how to understand QCA's notation system. Recall that QCA typically identifies combinations of factors that can be associated with the outcome. In the notation, these are connected by using the *, which represents AND. For example, the solution "Med*Equip" can be translated to read 'the provision of essential medicines AND equipment' combines to reduce MMR by at least 30%. Recall also that QCA typically yields more than one solution (e.g. various sufficient pathways), which can be connected using the +, which represents OR. For example, the solutions "Med*Equip" + "VHT*Staff" can

[24] However, there was one instance in which training formed part of a pathway, in Kicwamba III. However, we considered this pathway too unique to justify inclusion in the final pathways.

read the provision of 'essential medicines AND equipment' OR 'the provision of VHTs AND extra staff' both are associated with reducing MMR by at least 30%. Finally, the tilde (~) indicates the absence of a condition, e.g. ~medicines*equipment would read "the absence of medicines, combined with equipment…."

Step three involved running tests for necessity and sufficiency in both the government and private facility models after all of the above-listed modifications in place. We first examined the tests for necessity. Because necessary factors are 'always' present in all cases that share a certain outcome, there is a risk that they can be excluded by the Boolean minimization process as redundant, i.e. dropped from the solution.[25] Necessity tests therefore should be conducted separately from tests for sufficiency. We obtained the following results:

	Consistency
staff.5	0.511111
mento.p	0.511111
equip.27	0.688889
med.61	0.666667
infra1.1	0.133333
vht.1	0.466667
vbod.1	0.533333
ambu	0.021778

Pure necessity would involve a consistency score of 1 (with a .9 indicating "almost always necessary," but none of the above listed factors come close to meeting the threshold for necessity. An independent QCA reviewer, who had developed his own script to run more complex texts for necessity obtained the following results:

	Consistency
med.61+~vht.1	0.911111

Translated into English, this reads "either the provision of essential medicines OR the absence of VHTs is *almost always necessary* to reduce MMR by at least 30%."[26] While providing medicines certainly does make substantive sense, not providing VHTs as a necessary course of action does not. Therefore, we did not identify any necessary conditions in the model. Although medicines do play a prominent role in the outcome, they cannot be said to meet the criteria as a necessary condition.

For the private facility model, we also did not find any necessary factors when testing for necessity.
Analysis of Necessary Conditions

Conditions tested:

	Consistency	Coverage
mento.p	0.125000	1.000000
ambu.p	0.250000	1.000000
vbir.p	0.625000	0.714286
vbod.1	0.625000	0.714286
vht.1	0.125000	0.500000
med.61	0.875000	0.777778
equip.27	0.500000	0.800000

Here medicines come very close to meeting the criteria for "almost always necessary." However, we are obliged to point out that this is based upon only 8 health facilities and may have limited application elsewhere.

We then ran confirmatory tests for sufficiency.[27] The model produced solution findings that identify the various pathways, and their corresponding coverage and consistency scores. For the government model, our initial findings were as follows (we re-ordered the findings for ease of presentation):

[25] See Schneider and Wagemann (2012), pp. 221-225.
[26] See Rubinson (2015), "Review of Saving Mothers, Giving Life Phase 1 in Uganda: Cost Effectiveness Analysis," internal report, and SoCha (2015), "Response to Review of SMGL CEA," internal report.
[27] For the user, follow these instructions: For the government model, we followed these instructions:

Presentation of QCA Outputs				
Solution #	Solution Set	Raw Coverage	Unique Coverage	Consistency
1	vbod.1*~equip.27*~mento.p	0.133333	0.044444	1
2	vbod.1*~med.61*~staff.5	0.111111	0	1
3	vbod.1*~vht.1*~staff.5	0.155556	0	1
4	vbod.1*mento.p*~ambu.p	0.088889	0	1
5	equip.27*med.61*~staff.5*~vbod.1	0.133333	0.044444	1
6	equip.27*med.61*~mento.p*~vbod.1	0.111111	0	1
7	equip.27*med.61*~staff.5*~vht.1	0.133333	0	1
8	equip.27*ambu.p*~mento.p	0.133333	0.044444	1
9	equip.27*staff.5*~med.61	0.155556	0.155556	1
10	vht.1*staff.5*ambu.p	0.133333	0.022222	1
11	vht.1*staff.5*med.61	0.177778	0.022222	1
12	med.61*mento.p*vbod.1	0.222222	0.111111	1
13	med.61*mento.p*infra1.1	0.111111	0.066667	1
	Solution Coverage:	0.955556		
	Solution Consistency	1		
Note: vbod = Boda vouchers; equip= equipment; mento=mentoring; med=essential medicines; staff=hiring of additional staff; ambu=ambulances; vht=village health teams; infra=infrastructure				

In the raw QCA-readout form presented above, the government facility solution set is comprised of 13 different pathways. Overall solution consistency as well as individual solution consistencies all score 1 (perfect consistency), but this is because QCA dropped three facilities from the analysis that were in contradiction with each other (see above regarding Nyamabuga HC III). As a result of these dropped facilities, solution coverage decreased. We could have reversed this relationship, i.e. set a lower consistency threshold (e.g. .6) to keep the contradictions in, which would've increased our coverage score to 1, but decreased our consistency score to .978. Either is acceptable, but we make this point to highlight the inherent tension between consistency and coverage. Note also that five solutions have no unique coverage – which means that other solutions can also explain the facilities included in them. Some argue that these "zero unique coverage" solutions can be dropped. However, it would be a mistake to drop all of these without first confirming that the facilities on each zero unique coverage score pathway are not explained by another zero unique coverage score pathway (as this would eliminate some facilities from the explanation). In fact, several facilities only appear on zero unique coverage solutions. Because we don't have any strong theoretical reasons to drop any of them, we will leave them all in at this time.

---

1. Open fsQCA
2. Under File, Open Data - open the "Confirm Model 1 Government Facilities.csv" file
3. Under Analyze, chose "Crisp Truth Table Algorithm"
4. Under select variables, set mmr.3 (Our 30% reduction to MMR outcome) to Outcome.
5. The add Staff.5 (staffing), mento.p (mentoring), equip.27 (equipment), med.61 (medicines), infra1.1 (infrastructure), vht.1 (VHTs), vbod.1 (Boda vouchers) and ambu.p (ambulances). It is good to click show solution cases in output.
6. Hit Run
7. Review the truth table
8. Go to Edit - select "Delete and Code", use the default setting, which are:
    a. Delete rows with number less than: 1 (i.e. delete those truth table rows that have no entries)
    b. Set mmr.3 to 1 for rows with consistency >: .8 (this is the best practice setting for consistency cut offs)
9. Hit Standard Analyses
10. For the first round of PI Chart, hit "Mark All" then "OK"
11. Do this for the second.
12. For the Intermediate Solution (which is what we went with), select all factors as present (we expect the presence of these factors to contribute to a reduction in MMR; not their absence – we'll say more below)
13. Select "Mark All" and "OK" again.

A raw interpretation of many of these findings leads to two types of problems: a. nonsensical results, and b. potentially dangerous results. In regards to the first problem, the presence of several "non-presence" terms is problematic as policy recommendations. For example, solution 1 (*vbod.1*~equip.27*~mento.p) reads: "implementing Boda vouchers, combined with NOT providing equipment AND NOT mentoring staff is associated with reducing MMR by at least 30%." How can NOT providing resources be part of a program design, let alone explain why MMR reduced? We will discuss this in more detail below.

The second type of problem is more serious. Recall that the three delays framework seeks to reduce delays in seeking services, delays in reaching services and delays in receiving quality care. In practice, this involves providing resources to increase demand for facility-based births, and then at the same time provide additional human and material resources to meet that increase in demand within the facility to provide quality care. Although some of the solutions above suggest that Phase 1 was able to achieve success without providing all of these resources, they do not suggest that Phase 2 should consider addressing only supply side OR demand side concerns. We also discuss this in more detail below. For these reasons, we must adjust the QCA solution set to the realities of the SMGL context.

The Benevolent Assumption

To solve these issues, we make one crucial assumption: *all SMGL-activities we modeled either have a positive or, at worst, benign effect on MMR. Conversely, SMGL-activities cannot have a negative effect on reducing MMR.* In making this assumption, we carefully discussed each activity in detail with implementing staff to understand how the provision of each of these additional resources may lead to "unintended consequences" of doing harm. The only potential concern raised involved the provision of infrastructure, in which the renovation of a maternity ward resulted in its temporary closure, thus preventing mothers with complications from receiving vital EmONC during that time period. We investigated this possibility by looking at the timing of all maternal deaths in facilities that received maternity ward renovations. In all instances, none of the deaths occurred during the renovation periods, and thus we excluded this as a possibility. As such, we approach all SMGL activities with a "benevolent" assumption, i.e. doing more does not do any harm to that facility, and, conversely, not doing something cannot have a positive effect.

From QCA Readouts to Policy Recommendations

With the benevolent assumption in place, we were able to make four decisions that allowed us to resolve the two issues above associated with the QCA readouts and turn them into decipherable recommendations in line with the SMGL theory of change. Specifically, we:

1. Dropped negated conditions from the solutions
2. Combined solutions through factoring to create program pathways
3. Turned "ORs" into "ANDs" to widen the pathways
4. Where necessary, added more resources to ensure both supply side and demand side concerns were met.

We discuss each step in turn.

Dropping Negated Conditions from Solutions

Because we have assumed that the presence of a SMGL activity potentially contributed to positive changes in MMR, but the absence of a SMGL activity does not, we decided to drop the non-presence, or "negated" (those that begin with a ~) factors from the solutions. For scholars of QCA, this is a controversial decision, as the method rests upon "configurational" thinking in which variables must be understood in terms of their association/intersection with each other and not, as in statistics, in which each variable is considered in isolation of (or interaction with) each other. As such, dropping factors from a given configuration is nothing to be taken lightly. On the other hand, QCA advocates have recognized that the results of the software are vulnerable to "Type I" errors, i.e. those that identify configurations of factors as causal when they in fact just results of the way the data is structured.[28] To be sure, this possibility is there, but this is precisely why QCA-

[28] See, for example, Lucas, S. and Szatrowski, A. (2014), "Qualitative comparative analysis in critical perspective," Sociological Methodology 44: 1-79.

users must carefully reflect on the software results to ensure that the mechanisms identified are plausible and confirmed through qualitative investigation. As mentioned above, we did not find evidence to suggest that "not doing something" helps explain a reduction in MMR.

However, we did find another interpretation for understanding how non-presence factors combine with present factors to reduce MMR: assign contextual importance to non-presence factors as essential components of the solution. Here, solution 1 (*vbod.1*~equip.27*~mento.p) would read:

"implementing Boda vouchers in facilities where SMGL did NOT provide equipment AND did NOT mentor staff was still associated with reducing MMR by at least 30%."

What is important to point out here is that in many cases, the facilities on this solution set also did not receive many other types of SMGL activities (e.g. ~ambulances). Yet the algorithm used by the fsQCA software determined that these non-presence factors were not relevant, i.e. there was evidence to suggest that their presence or non-presence didn't make a difference when achieving the outcome. By contrast, the solution provided above suggests that only those particular non-presence functions, i.e. equipment and mentoring, were relevant contexts for reducing MMR (while the other non-presence functions weren't).

While this makes scholarly sense when explaining how SMGL was successful in the past, it does not make programmatic sense for predicting how SMGL can be successful in the future. All the various non-presence contexts identified by QCA in Phase 1 are not contextual factors in the sense that they exist independent of SMGL (i.e. they are not objective); instead, they must be seen as "emergent" in that they only can be conceptualized as a context when SMGL is being implemented. In fact, from a program design perspective (which is the one we adopt), it is irrelevant to the future design of SMGL to identify which "non-present factors" QCA has identified as relevant or irrelevant, because, from an implementer point of view, neither set of non-presence factors will be implemented.

Furthermore, it is also useful to ask, "But what does QCA suggest will happen if we do in fact provide Boda vouchers in combination with the provision of equipment and mentoring of staff?." The first answer is that we wouldn't see this configuration as a solution term. If QCA had that information – additional evidence that "Boda vouchers in combination with the provision of equipment and mentoring of staff" leads to success – then both the mentoring/~mentoring and equipment/~equipment would drop out as irrelevant, leaving only Boda vouchers. The second answer is connected to our benevolent assumption: "If we added those additional resources, at worst, nothing would happen to MMR (although we would waste resources), but it certainly would NOT cause MMR to decline." As such, we see no reason to include non-presence conditions in the set of recommended program pathways moving forward, as doing so does not add value to implementation decision making but does create potential confusion.[29] Before moving forward, we must however point out that if the goal is retrospectively explain how some Phase 1 facilities reduced MMR one way while others did it differently, re-introducing non-presence functions back into the explanation would provide a more comprehensive explanation.

Below we re-present the QCA solution set with the non-presence terms excluded:

Solutions after Non-presence Terms are Removed	
Solution #	Solution Set
1,2,3	Boda Vouchers
4	Boda Vouchers*Mentoring

[29] Although we discount non-presence factors here, oftentimes non-presence conditions do have important explanatory value. For example, Befani (2012) found that agricultural programs that offer farmers too many risk taking services (e.g. adopt new technologies, access to micro credit, etc.) during times of economic downturn can actually do more harm to the farmer than good. There, not doing part of the intervention certainly did matter for the outcome. We ourselves recently found that offering inkind grants to rural fishing communities in South Sulawesi, Indonesia actually undermined local business strategies in certain contexts. Even so, we maintain that theoretical sense must be upheld throughout any QCA analysis, and since including nonpresence terms violates this principle (and is potentially dangerous to mothers), we retain the right uphold substantive interpretation over algorithm orthodoxy.

5,6,7	Equipment*Medicines
8	Equipment*Ambulances
9	Equipment*Staff
10	VHTs*Staff*Ambulances
11	VHTs*Staff*Medicines
12	Medicines*Mentoring*Buda Vouchers
13	Medicines*Mentoring*Infrastructure

Dropping non-presence conditions then allows us to also combine solutions that say the same thing in terms "active ingredients," i.e. solutions 1-3 say the same thing, as do solutions 5-7.

Factoring

Although the above-listed modifications improve the coherency of the original QCA solutions, we are extremely reluctant to recommend them as program pathways for a Phase 2 SMGL. The reason why is because they violate SMGL's "3-delays theory of change." This is not surprising, as the algorithm QCA uses does not have the ability to identify how different pathways relate to each other, and is unable to detect how each solution may form inter-connected parts of a wider system. This is an extremely important point to bear in mind when bringing QCA to exercises like this, as the maternal health care system needs to be understood precisely as a system and not in terms of singular solutions.

Again using the first solution – only providing Boda Vouchers – as an example, it is useful to imagine if this solution was to be used as the basis for a new program design, i.e. only implementing Boda vouchers but forgoing additional support for human and material resources. The three delays framework tells us this would be irresponsible and irrational because:

1. Only providing Boda vouchers certainly not would address the needs of the maternity ward in a way that would reduce MMR in the facilities (as no within facility support would be provided).
2. It would, on the other hand, probably lead to increases in MMR, as the Boda Voucher program would create a steep rise in facility-based births, and overload the maternity wards without providing additional resources to address the increased demand.

Thus, we cannot isolate the role of Boda vouchers from other aspects of the intervention, and we are required to adapt the algorithm's findings to the context by combining pathways so that both supply side and demand side concerns are addressed. Yet in doing so, we need to retain the integrity of the original solution as much as possible and must also apply a logically consistent approach to combining these solutions. Fortunately, QCA offers a method for combing solutions, known as factoring. Factoring relies on rules of Boolean algebra to piece the above solutions together. For example:

Equipment*Medicines + Equipment*Ambulances ((Equipment AND Medicines OR Equipment AND Ambulances) = Equipment*(Medicines + Ambulances)

(Equipment, combined with *either* Medicines OR Ambulances, can successfully reduce MMR, although both are not needed).

Applied to all of the above solutions, the factored solutions are as follows:

Factoring as a Means of Combining Solutions		
Solution #	Solution Set (w/o non-presence factors)	Factored Solutions
1,2,3	Boda Vouchers	Boda Vouchers*Mentoring
4	Boda Vouchers*Mentoring	
5,6,7	Equipment*Medicines	Equipment * (Medicines + Ambulances + Staff)

8	Equipment*Ambulances	
9	Equipment*Staff	
10	VHTs*Staff*Ambulances	VHTs * Staff * (Ambulances + Medicines)
11	VHTs*Staff*Medicines	
12	Medicines*Mentoring*Boda Vouchers	Medicines * Mentoring * (Boda Vouchers + Infrastructure)
13	Medicines*Mentoring*Infrastructure	

We now have four solutions for government facilities, three of which present policy makers a set of options, or choices, regarding which additional factors they would like to add to reduce MRR, e.g. the second factored solution offers three different options (Medicines, Ambulances OR Staff) to combine with the provision of equipment to reduce MMR.

Turning "ORs" into "ANDs" to widen the pathways

Factored solutions such as those formulated above logically lead to a subsequent question regarding which option should be chosen. For example, in the third factored solution, the policy maker faces the choice between adding Ambulances OR Medicines to her/his VHT*Staff package. Answers to these questions depend upon one's philosophical orientation to how causality works across contexts. For the sake of argument, a purely "positivist" view might interpret these options as open choices based upon some kind of rational preference, such as which one costs the least to implement. By contrast, a "critical realist" view might instead interpret the choice as entirely context driven, e.g. that the choice of combining equipment with medicines or ambulances is determined by the needs of the context in which the intervention is administered. Our own views are more in line with the latter, and thus the answer to the question of which to choose is found through a more detailed analysis to define which contexts benefit more from ambulance vs. staffing vs. medicine type solutions (done of course in combination with equipment).[30]

Although we had a considerable amount of data to identify the applicability of these pathways to the wider context, we were not able to collect enough contextual information to answer this question at the lower levels. We therefore strongly recommend that implementers, at least at the beginning of the project, refrain from making this decision and instead implement all the choices associated with each factored solution, e.g. implement equipment in combination with ambulances, medicines and additional staff. In QCA-speak, we are recommending to turn ORs into ANDs, as it increases the chances that the solution will address the needs of the context, even if this may involve superfluous activities that are not required to reduce MMR to a similar level. The only exception to this method involves the use of infrastructure. Due to the specialized nature of infrastructure renovations, we do not recommend adding it to the rest of the package "to be on the safe side." Rather, infrastructure should be understood as a unique context variable based upon objective need. When this need is there, then we are advocating that the rest of the package identified in this solution also be implemented.

What are the effects of this decision on QCA-logic? From a retrospective view, this "widening of pathways" in fact has the exact opposite effect. As QCA-solutions add more and more factors connected by the conjunction AND, they refer to cases with more and more precision, or, put in QCA-speak, these solutions become subsets of much wider solutions. For example, the expression Ambulances AND Equipment can be understood as a subset of the superset Ambulances. Adding Equipment implicates a smaller number of facilities than those that received Ambulances in general. As such, solutions with more factors refer to a fewer number of facilities. This relationship is reversed when turned into a prospective policy recommendation. Here, adding ambulances and staffing and medicines to the equipment package yields a higher probability that MMR will be reduced by at least 30% than if SMGL only offered equipment. Again, recall that the benevolent assumption implies that adding these resources will not lead to any negative consequences in terms of outcome; at worst, they will be a

[30] For a discussion of how QCA, as a method, is compatible with both positivist and critical realist views, see Gerrits, L. and Verweij, S. (2013). Critical Realism as a Meta-Framework for Understanding the Relationships between Complexity and Qualitative Comparative Analysis. Journal of Critical Realism, 12 (2), 166-182.

waste of resources but QCA has already shown that their presence, in combination with equipment, can be associated with a higher reduction to MMR. Yet these solutions also can no longer be said to be direct products of QCA minimization. Although the terms we factored together were products of that minimization, we've now expanded those solutions by combining them with others. It is here we must acknowledge that we have made additions to the various solutions based upon substantive and theoretical concerns that sit outside of QCA-logic.

Compliance with the SMGL Theory of Change

Our last step was to ensure that the pathways we recommended conformed to the SMGL theory of change. This required us to make additional modifications to the solutions above. Presented below, these are:

1. Boda Vouchers * Mentoring
2. Equipment * Medicines * Ambulances * Staff
3. VHTs * Staff * Ambulances * Medicines
4. Medicines * Mentoring * Boda Vouchers * Infrastructure

To distinguish between the original QCA solutions and the ones we have derived here, we now use the term "pathways" to refer to the modified QCA solutions. In the four pathways, we find that pathways two to four all satisfy the requirements of the three delays framework, i.e. they consist of some kind of combination of providing resources to increase demand as well as provide both human and material resources to improve the quality of supply. However, the first pathways – Boda Vouchers AND Mentoring – does not meet this requirement because it lacks the provision of any additional material support, i.e. Equipment and/or Medicines.

We failed to find any QCA-related justification for expanding this pathway to include equipment or medicines. A review of the test for necessity scores for each reveals that they are practically identical in government facilities, and medicines score much higher in private facilities. Moving forward, the decision to provide either medicines or equipment (or both) should be driven by a context-specific understanding of what each facility needs, combined with a greater understanding of how both type of support interacts with the Boda Voucher program and the content of the mentoring instruction. However, we did pose a hypothetical question to both implementing partners, i.e. which do you regard as more valuable? In general, SMGL's ability to avoid stock-outs of essential medicines was the preferred choice. Therefore, we chose to add this provision to Pathway 1.

Finally, Pathways 1 and 2 include Boda Vouchers but do not include VHTs. Recall that Boda Vouchers were delivered via VHTs (i.e. Boda Vouchers are a subset of VHTs). Chapter Two discusses how this could occur at the facility catchment level under QCA. Although it is possible, these results still assume that some kind of VHT program is still in place, even if an individual facility-catchment may benefit from the vouchers without the presence of VHTs. As such, we include VHTs on these two pathways, and recognize that this isn't a reflection of QCA logic, but is a reflection of program logic.

The final four pathways for government facilities are then finalized as follows:

Recombined Government Pathways, in line with the Three Delays Framework				
	Raw	Unique	Facilities found only on this pathway	Facilities that also travelled other pathways
Pathway 1: VHTs * Boda Vouchers * Mentoring * Medicines[1]	0.3171	0.1463	(6) Bigodi HC III, Biguli HC II, Kagote HC III, Kasusu HC III, KISOJO HC III, KISOMORO HC III	(7) Bufunjo HC III, BUTUNDUZI HC III, Bwizi HC III, Kasenda HC III, Kigoyera HC II, Mahyoro Gvt HC III, Nyabbani HC III
Pathway 2: (Infrastructure)[2] VHTs[1] * Boda Vouchers * Mentoring * Medicines	0.3659	0.1951	(10) Kakumiro HC IV, KAMWENGE HC III, KATOOKE HC III, KIGARALE HC III, KYARUSOZI HC IV, Rukunyu HC IV, Rwamwanja HC III, Kigoyera HC II, Rwimi HC III	(6) Bufunjo HC III, BUTUNDUZI HC III, Bwizi HC III, Kakindo HC IV, Kibaale HC IV, Nyabbani HC III
Pathway 3: Equipment * Ambulances * Staff * Medicines	0.4146	0.2927	(12) Kabende HC III, Kibiito HC IV, Kicheche HC III, Kicwamba HC III (Kabarole), KIRYANGA HC III, Kisiita HC III, Kiyombya HC III, Kyabasaija HC III, KYEBANDO HC III GOVT, Ntara HC IV, Rugashari HC III, Rwagimba HC III	(5) Kigoyera HC II, Mahyoro Gvt HC III, MUGARAMA HC III, Nkooko HC III, Nyabbani HC III

Pathway 4: VHTs[1] * Staff * Ambulances * Medicines	0.2195	0.0732	(3) Kibaale Kasambya HC III GOVT, MABAALE HC III, Mpeefu Kasojjo/ Mpeefu B HC III	(6) Bufunjo HC III, Kakindo HC IV, Kasenda HC III, Kibaale HC IV, MUGARAMA HC III, Nkooko HC III
[1]Medicines and VHTs are not the result of combining pathways but are included based upon IP preference and program logic, respectively. [2]This Pathway does not require infrastructure, but for those facilities that receive infrastructure upgrades, the other elements should be included				

Private Facilities

We ran a similar QCA model for private facility catchment areas, which included Boda Vouchers, Equipment, Medicines, VHTs and a private facility-only activity, Delivery Vouchers. We excluded those activities that were only offered to government facilities, i.e. infrastructure, ambulances, mentoring and staffing. There were eight facilities in this model that had achieved success. The results were as follows:

	raw coverage	unique coverage	consistency
	---	---	---
~vbod.1*~equip.27	0.250000	0.000000	1.000000
~vbir.p*~equip.27	0.250000	0.000000	1.000000
vht.1*med.61*equip.27	0.125000	0.125000	1.000000
vbir.p*vbod.1*~vht.1*med.61	0.625000	0.625000	1.000000

solution coverage: 1.000000; solution consistency: 1.000000

Here, the results are slightly more complicated but we applied the same principles as above to improve their usability as program pathways and readability as active ingredients. The first two solutions are problematic because they consist entirely of non-presence functions, i.e. doing nothing at all reduces MMR under SMGL. Moreover, these solutions both apply to the same group of facilities – Midas Torch and St. Edwards. We had visited Midas Torch as part of our field visits and held a lengthy discussion with their chief Medical Officer. The conversation revealed that, in fact, Midas Torch had received additional support from several other implementing partners (some of whom were affiliated with SMGL). We could not model this information systematically across all facilities (i.e. identify which received additional support from additional IPs), but this could explain the counter-intuitive results of the first two solutions presented above.

We examined the remaining two solutions in terms of their coverage scores. Although substantially larger than most in the government model, the smaller case size (8 facilities) means that the solutions applied to much fewer cases. In fact, solution 3 only applied to one facility – Kiko HC II. We therefore dropped solution 3 from the recommended pathway list on the grounds that it was too infrequent to be included.

This leaves solution four, which contains a non-presence factor, i.e. ~vhts. We followed the same logic as above and dropped this non presence factor, which resulted in Pathway 5. We also had to add VHTs, as Boda Vouchers assumes their existence:

Pathway 5: Delivery Vouchers * Boda Vouchers * Medicines
Facilities: MITANDI HC III, Yerya HC III, Padre Pio HC III, Kicwamba HC II, St. Martins Mabira HC III

Recombined Private Facilities Pathway, in line with the Three Delays Framework				
	Raw	Unique	Facilities found only on this pathway	Facilities that also travelled other pathways
Pathway 5: VHTs [1]* Boda Vouchers * Mentoring * Medicines	0.625	0.625	MITANDI HC III , Yerya HC III, Padre Pio HC III, Kicwamba HC II, St. Martins Mabira HC III	None
[1] VHTs are not the result of combining pathways but are included based upon program logic.				

Note the lack of human resource support in this pathway. Although the three delays framework would suggest human resource support be provided, SMGL phase 1 implementers did not offer these services. Instead, private facilities were expected to re-invest the revenue gains made from the delivery vouchers into hiring additional as well as mentoring their own delivery staff. Thus, Pathway 5 deviates from the other pathway format.

Annex 3: How We Modeled SMGL's Cost

Guiding Principles

A cost model is only as good as the assumptions and inputs upon which it rests. We made a concerted effort to ensure that the data we used was reliable, but we do recognize that the usefulness of our price, demand and input data can change as the context changes over time. As such, we have built it to be fully customizable and transparent in how it operates. The user can select which districts, or combination of districts, they would like to use and the model is instantly populated with demand data – population growth and delivery at facilities – to combine with our input ratios and prices to yield costs. The user can customize costs per district, per activity, disaggregate by facility type, etc. over five years, accounting for both startup and operations cost, broken down by activity. If the user disagrees with any of our cost estimates/ratios, salaries, etc.; they can change the amounts accordingly and these changes will carry through in the model. We layout 3 different scenarios based upon different population growth assumptions.

We adopted two main principles to help guide the costing analysis. First, we adopted the "expectant mother view" to costings. This approach starts with the view of the mother as she begins her journey through the Ugandan maternal health care system, and seeks to cost all services associated with that journey, from village health team to antenatal visits to delivery and the first 48 hours post-delivery. In doing so, we mapped all the associated organizational processes and sub processes, and further divided these according start up and operating costs. The result of these process maps were then used as the blue prints for the construction of the costing model. In total, we generated 19 maps that guide how we costed Health Facility and SMGL service implementation; and cross checked these with the implementing partners both in the field and in Kampala to ensure their accuracy. These process maps formed the blue prints of the costing model, and they are attached in Annex 5.

The second principle we adopted was to draw clear boundaries around what was to be costed, and limited our costing to only those services which SMGL sought to improve. We further refined the costing boundaries by posing the question "What does SMGL wish to take to scale?" to guide what was in and what was out. For example, imagine a community health worker encourages a pregnant mother to seek testing for malaria; she does so, and unfortunately tests positive. Where the SMGL costing stopped was at the referral process, and we did not continue to follow the wider costs associated with providing the mother with anti-malarials. The USG has other mechanisms to improve non-maternal health related service delivery from which the maternal mother could benefit, but costing these services was irrelevant to this exercise as SMGL is not seeking to take these services to scale. This principle also helped us when we come to the question of whether or not to cost the inputs of other donor programs. As SMGL was not seeking to take other donor programs to scale, understanding the amount they may have spent at SMGL-facilities is not relevant. On the other hand, because the SMGL program was designed to support and strengthen the government and private health care system, we also estimated baseline delivery costs, upon which SMGL activity costs were added.

Costing Methodology and Model Design

The cost of the activities in the maternal health care system are driven by the demand for services, which drives the quantity of services required, the resources or inputs consumed in the activities and the price of these activities. When costing a program, it is important to understand the flow of activities, the facilities involved in the activities, the range of inputs used and the price of these inputs. Many of Phase 1 activities were also demand driven, for instance the provision of transport vouchers. However, other activities depended on the program design and explicit decisions made by the program implementers to provide a specific level of support, for example the purchase of additional equipment for facilities and the upgrading of infrastructure. In these later instances, assumptions were developed that sought to replicate the kinds of decisions that were made in the SMGL Phase I implementation.

Even though we introduced a great deal of complexity into the model with the process maps, our approach always came back to the simple costing formula: Cost=Demand x Inputs x Price. This was in turn divided into two different conceptual methodologies. Broadly speaking, the two approaches were used:

v. *Activity based costing* for those components that entail implementing a process such as delivering a service or providing support. These processes can be once-off in nature (i.e. set-up activities) or they can be continuous in nature (i.e. service delivery activities).

vi. *Capital costing* for those components that involve building facilities and supplying equipment such as vehicles and medical equipment.

The methodologies used to cost the maternal health care system were:

vii. *Institution based costing* for staff and equipment:

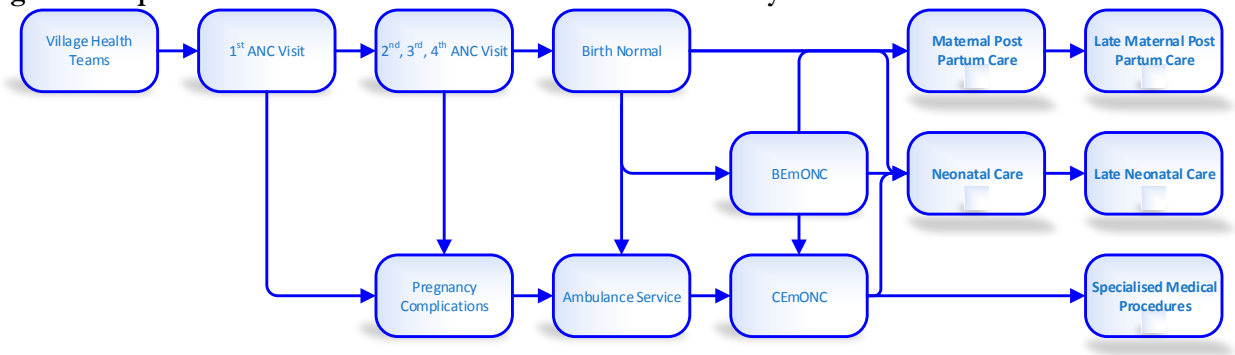
- average staffing levels for the different health care facilities are used to calculate the cost of personnel employed at health facilities to deliver maternal health services; and.
- the amortization cost of equipment the health facilities required to deliver maternal health care services, adjusted by a factor that reflects the undersupply of such equipment.

viii. *Cost of procedures* for medicines and supplies used within the maternal health care system.

Design of the SMGL Costing Model

As noted above, the "expectant mother view" to the maternal health care system informs the design of the costing model where appropriate. Understood in broad terms, the main elements of the expectant mother view can be summarized as follows:

Figure 3: Expectant mother's view of the maternal health care system



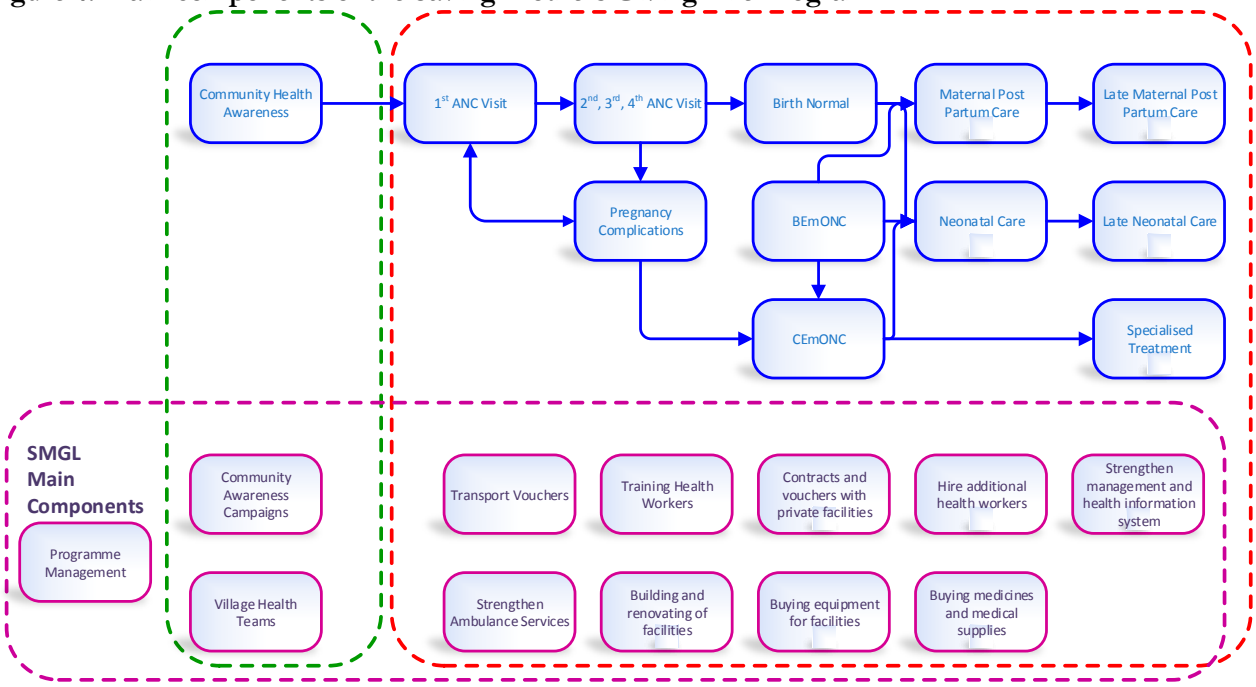
This summary should be understood as broad baseline processes that were already in place prior to SMGL's commencement (i.e. the Ugandan health care system costs). The associated costs of each of these elements are highly influenced by varying levels of demand and uptake, especially costs that involve medicines and supplies. These baseline services were built around the following facilities:

- Ambulances
- Level III health facilities
- Level IV health facilities
- District hospitals (which includes regional hospitals)

Layering SMGL Phase 1 onto the Expectant Mother View

The "expectant mother view" is less apparent in the sections of the model that cost the SMGL program. This is because Phase 1 activities were designed to strengthen and support the Ugandan maternal health care system and not replicate it. Therefore, we built the master process map by linking Phase 1 activities to the Expectant Mother view in the figure below:

Figure 4: Main components of the Saving Mothers Giving Life Program



Most of the activities in Phase 1 were meant to support antenatal visits, delivery and the first 48 hours post-delivery. These are highlighted by the components in the red-dotted rectangle in the above figure.

Structure of Cost Categories

As discussed in the previous chapter, SMGL Phase 1 consisted of 24 separate activities, and we combined many of these to identify 10 factors we used for the QCA model. Two remaining cost categories were excluded as they could not be assigned as health facility level properties: systems-strengthening activities and program management. However, for the cost model, we were able to incorporate these two categories into the analysis. We followed the same grouping procedures for each model; e.g. all the training related activities were grouped into "Training Health Workers" – and then the detail in the costing sets out the different kinds of training that was provided. We present the cost results in three formats. Firstly, the cost information is presented by type of facility and by Phase 1 activity as follows:

Figure 3: Presentation of the costing results by type of facility and SMGL component – Set Up Costs

Costing Summary

	Baseline	Year 1	Year 2	Year 3
	Baseline	SMGL	SMGL	SMGL
Health Care Costs - funded by the Ugandan Government - All Values are in USD				
Operational Costs	**4,749,292**	**5,861,515**	**6,365,945**	**6,908,536**
Ambulances	67,800	70,663	72,968	75,630
Level III Health Facilities	2,424,226	3,050,551	3,360,590	3,850,967
Level IV Health Facilities	1,211,816	1,413,285	1,512,881	1,622,268
Hospitals	1,045,450	1,327,016	1,419,506	1,359,671
Replacement of Equipment	**316,228**	**333,937**	**351,969**	**371,679**
Level III Health Facilities	135,359	142,939	150,658	159,094
Level IV Health Facilities	110,033	116,194	122,469	129,327
Hospitals	70,836	74,803	78,843	83,258
Costs of the SMGL Program	**All Values are in USD**			
Total Annual SMGL Costs	**-**	**10,513,086**	**7,006,744**	**7,341,369**
Set-up / Once-off Costs	**-**	**4,761,823**	**918,408**	**870,131**
SMGL management	-	-	-	-
Awareness Campaigns	-	-	-	-
Village Health Teams	-	275,117	7,999	8,014
Transport Vouchers	-	24,931	1,383	1,386
Ambulances	-	903,950	-	-
Training	-	-	-	-
Hiring additional personnel	-	-	-	-
Private Facilities' Vouchers	-	-	-	-
Medicines purchased	-	-	-	-
Equipment purchased	-	1,224,318	202,586	213,931
Upgrades, Renovations and Builds	-	2,333,508	706,440	646,800
System Strengthening	-	-	-	-

Figure 4: Presentation of the costing results by type of facility and SMGL component – Operational Costs

Costing Summary - Continued				
	Baseline	Year 1	Year 2	Year 3
	Baseline	SMGL	SMGL	SMGL
Operational Costs	-	**5,751,263**	**6,088,336**	**6,471,238**
SMGL management	-	1,400,391	1,480,628	1,569,235
Awareness Campaigns	-	344,320	343,668	344,320
Village Health Teams	-	222,997	222,580	222,997
Transport Vouchers	-	218,481	218,067	218,481
Ambulances	-	96,304	83,664	86,384
Training	-	552,188	551,142	552,188
Hiring additional personnel	-	701,942	746,866	796,159
Private Facilities' Vouchers	-	788,366	839,331	893,482
Medicines purchased	-	1,244,730	1,421,185	1,606,448
Equipment purchased	-	-	-	-
Upgrades, Renovations and Builds	-	-	-	-
System Strengthening	-	181,545	181,204	181,545

The above listed figures come from the Costsum Pilots tab in the cost model, and have been divided into two figures for ease of presentation. The totals that are listed are for the four SMGL pilot districts, and assume 65,490 births (taken from DHMIS data). Years 1 through 3 assume moderate increases in the birthrate, and can be adjusted on the Scenarios tab of the model. Baseline year costs of course do not include SMGL costs, which come into effect in Year 1. Note the division of the SMGL component costs into the self-explanatory categories of "Setup / Once-off Costs" and "Operational Costs".

Secondly, the cost information for each type of health facilities is presented as follows:

Figure 5: Presentation of the costing results by type of health facility

HS3. Level III Health Facilities					
Item		Costing scenarios			
		Baseline	Year 1	Year 2	Year 3
Summary of costs		Baseline	SMGL	SMGL	SMGL
Normal Operational Costs *(excludes replacement of equipment)*	USD	**2,424,226**	**3,050,551**	**3,360,590**	**3,850,967**
Personnel	USD	1,222,465	1,259,139	1,339,724	1,428,146
Medicines and medical supplies	USD	722,494	1,285,306	1,515,718	1,916,715
Transport	USD	67,162	70,923	70,788	70,923
Other goods and services	USD	412,105	435,183	434,359	435,183
Replacement of equipment (based on depreciation)	USD	135,359	142,939	150,658	159,094

Note the item "Replacement of equipment" is not included in the "normal operational costs" of the facilities because currently there are no routine asset management systems in place to ensure that equipment within the health sector is replaced once it has served its useful life. These figures are also strongly dependent on what equipment is currently available and the age of this equipment – both of these variables would require a detailed audit of current equipment (which was not available when the costing model was being built).

Thirdly, the cost information for the SMGL components is presented as follows:

Figure 6: Presentation of the costing results by SMGL component

SMGL Programme Management

Item		Costing scenarios			
		Baseline	Year 1	Year 2	Year 3
Summary of costs		Baseline	SMGL	SMGL	SMGL
Set-up costs		-	-	-	-
Personnel					
Medicines and medical supplies					
Training					
Other goods and services					
Capital - medical equipment					
Capital - vehicles					
Capital - renovations and building					
Operational costs	**USD**	-	**1,400,391**	**1,480,628**	**1,569,235**
Personnel	**USD**	1,221,287	1,257,926	1,338,433	1,426,770
Medicines and medical supplies	**USD**				
Training	**USD**				
Repairs and maintenance	**USD**				
Other goods and services	**USD**	134,910	142,465	142,195	142,465

There are in fact 12 tabs under the SMGL cost model section, one for each activity in the model.

Annex 4 contains a user's manual, but other points worth noting: We assume a UGS/USD exchange rate of 2500 to 1, assume a 6% inflation rate, and present the full range of our assumptions on the GenAssumptions tab. Cost assumptions around build constructions, equipment and medicines/supplies delivered per complication are presented in the HS Build, HS Equip and Supplies tabs, respectively. Finally, complication ratio, salary and district demographic assumptions and data are presented in the Base/SMGL Ratios, Salaries, and District Data tabs, respectively.

Data Sources

We worked closely with Baylor and IDI to compile expenditure data, and turn it into cost data by associating it with their organizational structure and processes. We found this to be an extremely effective approach to what is typically a very sensitive subject – the budget. The goal was not to seek a historical recreation of all that was spent in Phase 1; instead, it was to collect enough information to develop cost ratios that could be applied elsewhere. We sought to either incorporate budget information directly in the model or develop assumptions that approximated the implementation experience of the partners. Once we effectively communicated that our purpose was in no way connected to audit or similar functions, the partners offered levels of cooperation and support beyond our expectations. In addition, we had access to the raw expenditure data collected by the Futures Group for the *Expenditure Analysis for Saving Mothers, Giving Life,* which provided detailed expenditure and price information which was used in the appropriate places in the costing model.

The following table outlines the main data sources used on the various SMGL Costing Model spreadsheets:

Table 1: Main data sources for the SMGL Costing Model spread sheets

Sheet in Model	Data	Data sources
GenAssumptions	Wide range of price information	• Baylor Consolidated Budgets • Futures Group Study
Scenarios	Assumptions about birth rates and infant mortality rates	• Uganda Demographic and Health Survey (2011)
District Data	Population per district and number of health facilities per	• Uganda Bureau of Statistics; National Population and Housing

	district	Census 2014 (Provisional Results) • List of Health Facilities from the Department of Health Note that the district data drives most of the demand related calculations in the model
Salaries	Salary grades and salary levels	• Futures Group Study
Base Ratios **SMGL Ratios**	Complication ratios which show the flow of mothers through the system	• Department of Health Information System data (drawn and analysed by SoCha) • SMGL CEA Survey (conducted November 2014)
HF Level III **HF Level IV** **HF Hospitals**	Number of staff per facility, equipment levels, construction requirements	• SMGL CEA Survey Salary information from the Salaries sheet and price information from GenAssumptions feed into these sheets
SMGL sheets	Descriptions of SMGL activities and processes	• *External Evaluation of Saving Mothers, Giving Life: Final Report* (October 2013) • Information from site visits and interviews
HS Build	Cost of construction works	• Futures Group Study • Baylor Consolidated Budgets The estimates for a lot of works used in the costing were made based on these data sources and Parker et al.
HS Equip	Prices of medical equipment	• Joint Medical Store: Catalogue and Price Indicator February 2014 • Supply Requisition Sheets from Baylor and IDI • Futures Group Study • Parker et al[66] • Dikshin Uganda (water tanks) • Various websites of retailers
Supplies	Cost of medical supplies used in procedures	• Original list from IDI, but modified in consultation with doctors working for Baylor and IDI This sheets feeds into the Base and SMGL ratios sheets)

Site Visits, Interviews and Collection of Additional Information

To build the Phase 1 and expect mother process maps, we visited Fort Port Regional Referral Hospital, two District Hospitals, 5 Level IV health facilities and several Level III health facilities across the four districts, as well as paid several visits to both IDI and Baylor's field and Kampala offices. Staff included program managers and medical specialists, as well budget managers from Baylor and procurement officials from IDI. These visits and interviews provided valuable process information for the building of the SMGL Costing Model, and also built contacts with managers and medical specialists who provided additional background information on SMGL Phase I, as well as participated in developing the facility equipment lists (HS Equip) and the costing of the medical procedures (Supplies). We also commissioned a SMGL survey of a sample of health facilities that formed part of the SMGL Phase I. This survey gathered information on facility staffing level norms, the availability of equipment and the purchasing of medicines and supplies that was used in developing the applicable assumptions in the SMGL Costing Model.

We applied significant effort to deriving accurate estimates for the demand, inputs and prices surrounding various complications and procedures during delivery. We first analyzed data from the POMS dataset. POMS is comprised of two subsets divided according to level of facility, i.e. subset 1 is from Level II and III facilities, while subset 2 is from Level IV and above. We primarily used subset 2, because it focused on monitoring maternal outcomes across 40,526 births individual births, 21,822 of which were during Phase 1. As with the challenge of standardizing health facility names, there were no standardized protocols for naming various complications and we faced an initial list of 306 complications. POMS also contained four separate variables for complications (Cset, obs_com, dir_com and M_comp1 & 2), which unfortunately didn't yield

[66] Parker, Cruddas, Rowe, Carter and Webster (2012). Tank costs for domestic rainwater harvesting in East Africa. Water Management. Volume 166 Issue WM10.

consist totals for the same complication across the four variables. We sat with medical officers from both implementing partners and distilled this list to 43 sets of complications. We then worked with the medical officers to identify which sets of procedures required corresponding medicines (and their units of dosage), medical supplies and equipment used, as well as the estimated labor inputs per complication in terms of critical personnel, e.g. medical officers, nurses, midwives, etc. This gave us our input data. We matched these inputs to price data drawn from official NMS and JMS procurement lists that contained quantities and costs for 223 items associated with maternal health and emergency obstetric care. The totals gave us the average unit cost of addressing each complication. Finally, we derived the demand data by standardizing the complication names in the dataset so we could derive complication rates as a percentage of total births in the dataset.

Limitations

Wherever possible, the SMGL Costing Model is based on data collected during the implementation of SMGL in Phase I. A wide range of data from the implementing partners and other consulting projects related to this phase of the SMGL program are used. In some instances, the data was obviously flawed in which case we did not use it – for instance the reported number of mother's attending a first ante-natal visit was reported to be 428 per cent the number of births – which is a very unlikely number.

In certain instances, there are gaps in the data and therefore a number of estimates and assumptions had to be made. In some cases the SMGL program work consisted of a wide variety of pieces of work, particularly in relation to repairs to infrastructure and building work. In these cases a typology of the activities was developed and the actual information used to populate it as far as possible, and then further assumptions were made to fill in the gaps. In some regards, the costing represents a middle ground or an average of what seems to have occurred in the SMGL Phase I – based on the available information.

In other instances, e.g. the supplies sheet (Supplies) and the equipment sheet (HS Equip), the costing is based on very precise price and demand information drawn from what appear to be good data sources supplied by the implementation partners and confirmed with medical specialists working for the partners. These sheets can support very detailed planning and support other activities such as procurement.

The model was designed to cost the components of the SMGL program and its possible rollout to additional districts using an 80/20 approach. This means that we aimed to do the first 20 percent of the effort that produces 80 percent of the estimated budget – given the time and budget constraints on the project. There is therefore a margin of error in the model that must be considered when interpreting its results – it is very difficult to ascertain this margin of error with any certainty in the absence of a clear counterfactual. Nevertheless, comparisons with the Futures Group expenditure analysis do suggest that the SMGL Costing Model produces costing results that are in the right order of magnitude when compared to the expenditure outcomes of SMGL Phase I.

Annex 4: Cost Model Instruction Manual

Basic Concepts for the User

This manual is intended to be brief. The intention is to direct the attention of the user to the model rather than distract them with words. Effort has been made to make the model simple and user friendly. The variables are described in ways that should make sense to anyone familiar with SMGL, and should be understandable to anyone who has a high level understanding of the activities and objectives of SMGL.

There are some critical rules that the user must observe. The model is not fool proof. If the model is not used as intended, errors are likely to occur. Care has been taken to ensure that if a user enters a number that causes a mathematical error (e.g. dividing by zero) the model will continue to function, except where an error is caused by the user entering a number that is not plausible. This section aims to make the user familiar with these rules. Section 2 provides a high level explanation of each worksheet.

Protected Sheets, hidden rows and secret numbers

The worksheets in the model are protected. The nature of the protection may vary between worksheets. What this means is that, in some worksheets the user can select protected cells, while in other worksheets the user will not be able to select protected cells. Most worksheets do not need a password to unprotect them.

In almost every sheet many rows have been hidden. This has been done to hide calculations or hide empty rows that are part of a template and are hidden to make the sheet easier to read. Users may unprotect sheets and then unhide rows to see the calculations if they are interested.

In many places in the model there are numbers that are in very light shades of grey that may be visible. These are used in formula (functions). They will not make any sense to someone who does not understand Excel functions (e.g. VLOOKUP, INDEX), but they are crucial to the functioning of the model and should not be changed.

Colour Formatting - Yellow versus Blue

Even though the model is protected and users will be prevented from changing cells they do not have the authority to change, it is still crucial that the following rule is complied with:

You cannot change values in cells that are this colour
You can change values in cells that are this colour

The above formatting **has not been** applied to the "Supplies" sheet. Changes to the "Supplies" sheet should only be made by medical experts and it has therefore been left non-user friendly.

Drop Down Menus

Drop down menus are used in most of the worksheets. Most cells with drop down menus have words in them. They are all shaded in blue, but some will change from Blue to White if the option "No" is chosen, as shown in the following example:

Support for Planning workshops with Maternal Health Committees				
Number of workshops per district/year	2	2	2	2
Number of participants per workshop	10	10	10	10
Travel pay for participants/per participant	Yes	Yes	Yes	Yes
Transport refund for participants	Yes	No	No	No

This has been done to emphasise to the user when certain costs are not factored into the costing (transport refunds for the second to fourth years in the above example).

If a cell with a drop down menu is left empty, the cell will become red to draw attention to it, as in the following example:

Support for Planning workshops with Maternal Health Committees

Number of workshops per district/year	2	2	2	2
Number of participants per workshop	10	10	10	10
Travel pay for participants/per participant	Yes			Yes
Transport refund for participants	Yes	Yes	Yes	Yes

If a cell with a drop down menu is left empty, the model will either use a default value to estimate costs, which may not be based on the correct assumptions, or an error will show up.

Example One: Standard Sheet Layout

Village Health Teams

Item	Costing scenarios			
	Baseline	Year 1	Year 2	Year 3
Summary of costs				
Set-up costs	**3 398 306**	**3 413 870**	**96 052**	**96 234**
Other goods and services	3 398 306	3 413 870	96 052	96 234
Operational costs	**2 423 813**	**2 557 483**	**2 552 709**	**2 557 483**
Other goods and services	2 423 813	2 557 483	2 552 709	2 557 483

A summary of costs calculated within the worksheet is located at the top of each worksheet.

Input and process assumptions				
No of pregnancies	1 745 637	1 745 637	1 745 637	1 745 637
Demand assumptions				
Number of pregnant women for each VHT	150	150	150	150
Number of VHTs required	11 638	11 638	11 638	11 638
% of VHTs that leave per year	3%	3%	3%	3%
Number of newly elected VHTs each year	349	349	349	349
Per cent of VHTs that attend recruitment sessions that are selected	50%	50%	50%	50%
VHT Coordinators	*if there are no coordinators, then enter 0 (zero) in both rows below*			
Number of VHTs per parish co-ordinator	7	7	7	7
Number of Parish co-ordinators per Sub-Co	6	6	6	6
Number of Parish Co-ordinators	1 663	1 663	1 663	1 663
Number of Sub-County co-ordinators	277	277	277	277
Do you want to cost set up or maintenance?	Set up	Set up	Maintenance	Maintenance

The demand variables relevant to what is costed in the worksheet will appear here.

Blue cells can be changed – but do not necessarily need to be.

Yellow cells show results of calculations, and choices.

Effort has been taken to make the variables self-explanatory.

Blue cells with words generally mean the that a drop down menu is available.

Example Two: Personnel Assumptions and Job Titles

Personnel assumptions

Company national program office	**No. per company**	**% of time**
Program Supervision and Support		
Executive Director - N	1	15%
Technical Advisor - N	1	100%
Program Coordinator - N	1	100%
Quality Improvement Manager - N	1	100%

These are Job Titles. Each has a different salary, which is provided in the "Salary" worksheet. These are drop down menus. Care has been taken to choose the correct Job Titles throughout the model, so these should not need to be changed.

Explanations of the Sheets

The following pages explain the model through the use of detailed pictures. The user is encouraged to open the model and play around first. Then, when a more detailed explanation is required, these diagrams can be used to provide more information. If you understand the colour rules explained above, then the worksheets should be self-explanatory.

The contents of the next pages are as follows:

- A description of the key modelling options in the "Scenarios" worksheet. This sheet is very significant in the model.
- An overview of the model, showing the flow of information between worksheets. The colours of the lines do not mean anything; different colours are used simply to make the various connections clear. The direction of the arrows shows the flow of information.
- Worksheets with key modelling parameters, which are "GenAssumptions", "Scenarios", "District Selectors".
- Costing of Ugandan Health Care – four worksheets are described, namely "HS Ambu", "HS Level III", "HS Level IV" and "HS Hospitals".
- Explanation of the data sheets in the model, which are the "Salaries" and "District Selectors" worksheets.
- Complications worksheets, of which there are two: "Comps Baseline" and "Comps SMGL". These are key worksheets in the model.
- Three worksheets that contain information about the prices and levels of equipment, building works and medical supplies, are discussed together. These worksheets are "HS Build", "HS Equip" and "Supplies".
- The worksheets with the SMGL activities are described over two pages due to the quantity.

Scenarios

click here to choose districts and regions

	Baseline	Year 1	Year 2	Year 3
	Baseline	SMGL	SMGL	Baseline
		Per cent change relative to previous year		
Population	8 715 800	10%	0%	0%
Population under scenario	8 715 800	9 587 380	9 587 380	9 587 380
Uganda birth rate (births per 1 000 people)	47.4	47.4	47.4	47.4
Infant mortality rate (deaths) per 1 000 births	79	79	79	79
Number of Live Births	413 129	454 442	454 442	454 442
Number of deaths at birth	32 637	35 901	35 901	35 901
Number of Births	445 766	490 343	490 343	490 343
Per cent of births delivered at health facilities	80%	80%	80%	80%
Number of deliveries at public health facilities	356 613	392 274	392 274	392 274
Number of Health Facilities				
Level II	359	359	359	359
Level III	275	275	275	275
Level IV	46	46	46	46
D Hospitals	47	47	47	47
R Hospitals	-	-	-	-
Number of districts	24	24	24	24
Show whether year is a Baseline year or an SMGL year	Baseline	SMGL	SMGL	Baseline
Maximum Number of Births per month per facility				
Level III	60	60	60	60
Level IV	150	150	150	150
Hospitals	350	350	350	350
Rank level of health facilities in order of preference for delivery from the mother's perspective				
Level III	First			
Level IV	Second			
Hospitals	Third			
Number of deliveries per level of health facility				
Level III	198 000	198 000	198 000	198 000
Level IV	82 800	82 800	82 800	82 800
Hospitals	75 813	111 474	111 474	111 474
Number of deliveries health system can't serve	-	-	-	-

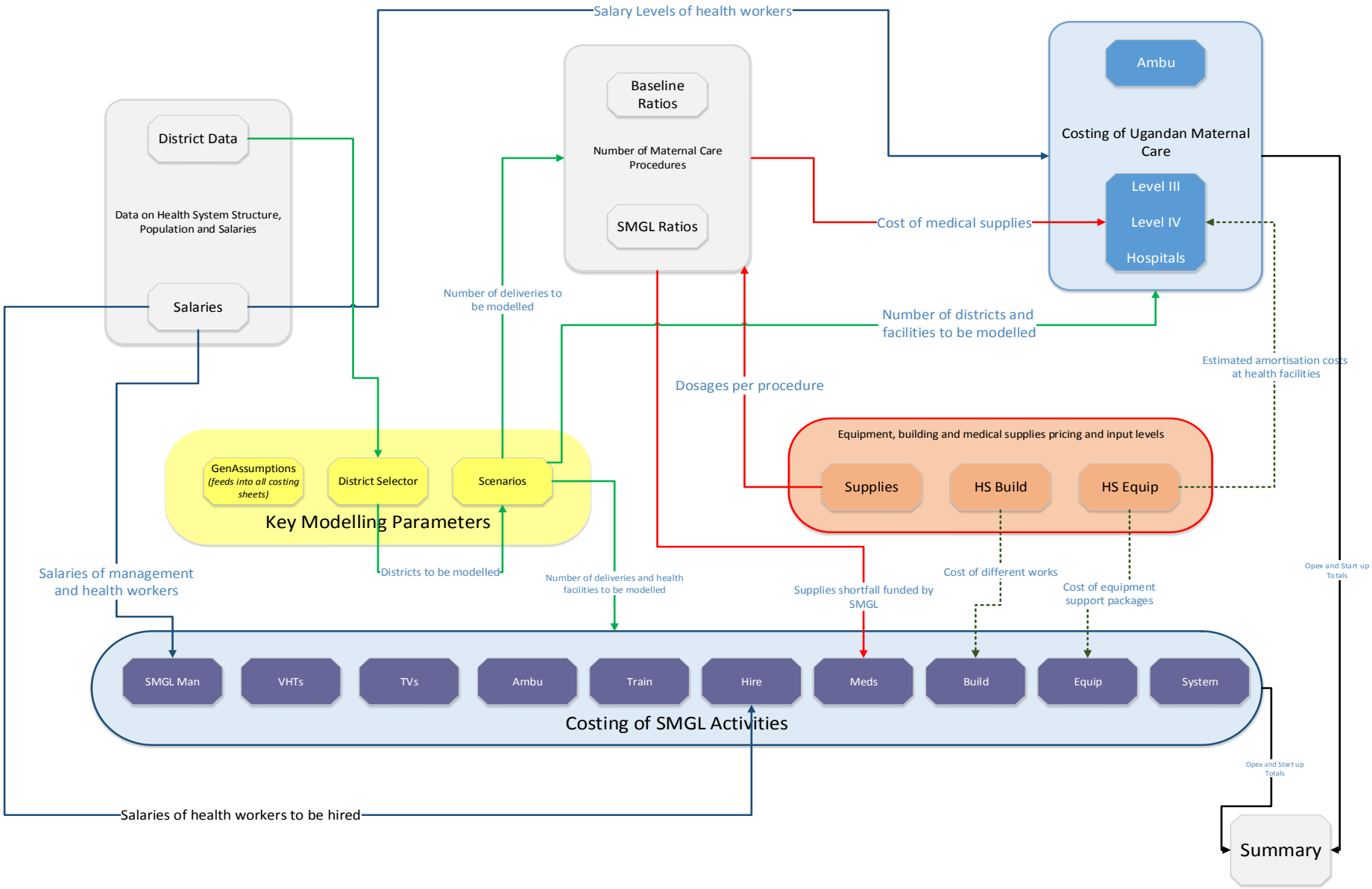
This is effectively a population growth rate.

Birth rate drives number of live births. Children lost through infant mortality rate is added to number of live births to calculate number of deliveries.

This is pulled through from "District Data" and depends on the selections in "District Selector". The rows indicated by the red bracket are the levels of demand costed.

This drop drown must be used to show whether the SMGL programme is being implemented in the year or not.

These two options determine how the deliveries cascade through the health system, and the results are shown below.


Salary Levels of health workers
District Data
Data on Health System Structure, Population and Salaries
Salaries
Baseline Ratios
Number of Maternal Care Procedures
SMGL Ratios
Ambu
Costing of Ugandan Maternal Care
Level III
Level IV
Hospitals
Cost of medical supplies
Number of deliveries to be modelled
Number of districts and facilities to be modelled
Estimated amortisation costs at health facilities
Dosages per procedure
Equipment, building and medical supplies pricing and input levels
Supplies
HS Build
HS Equip
GenAssumptions (feeds into all costing sheets)
District Selector
Scenarios
Key Modelling Parameters
Salaries of management and health workers
Districts to be modelled
Number of deliveries and health facilities to be modelled
Supplies shortfall funded by SMGL
Cost of different works
Cost of equipment support packages
Opex and Start up Totals
SMGL Man
VHTs
TVs
Ambu
Train
Hire
Meds
Build
Equip
System
Costing of SMGL Activities
Opex and Start up Totals
Salaries of health workers to be hired
Summary

Key Modelling Parameters

GenAssumptions *(feeds into all costing sheets)*

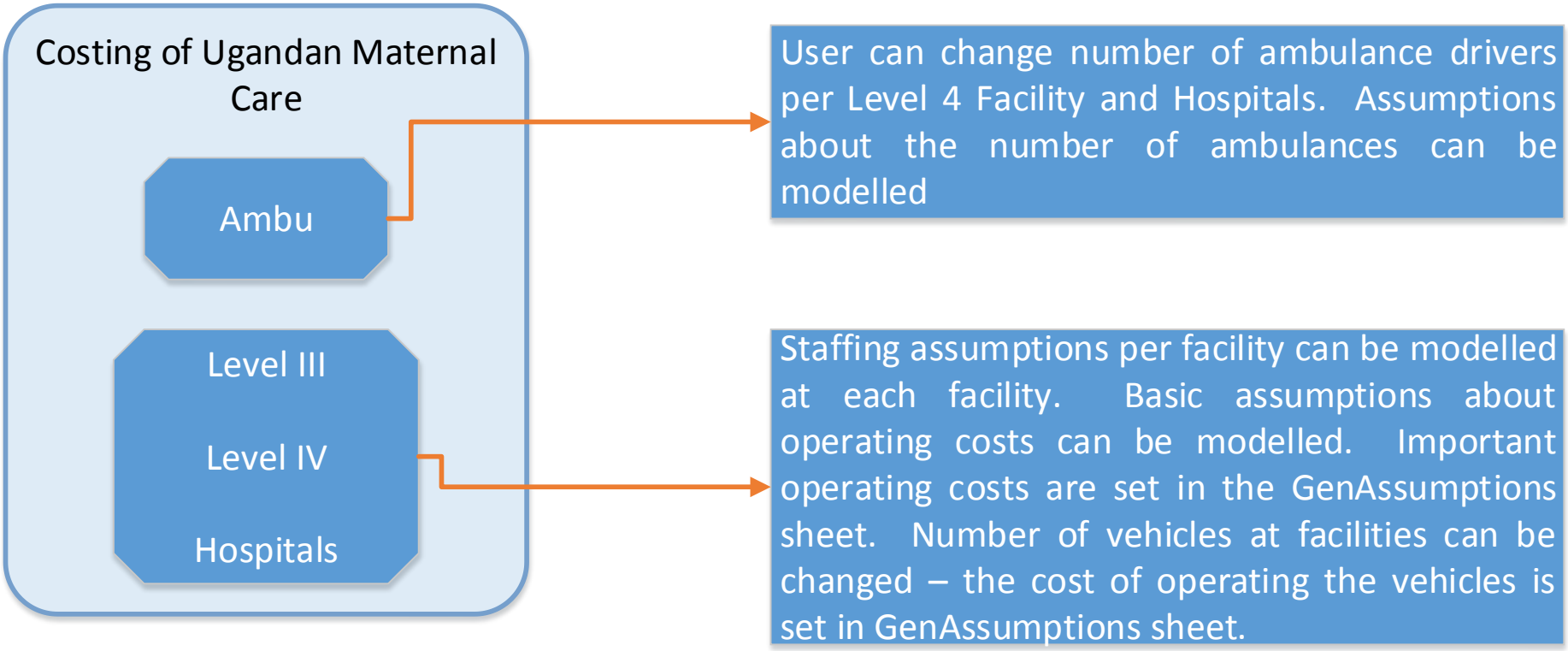
District Selector

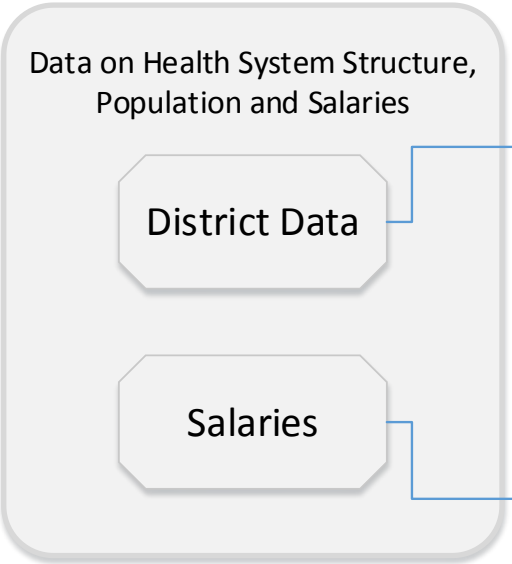
Scenarios

This sheet contains a long list of input price assumptions. Rates of inflation and exchange rate is also set here. Almost every sheet in which costs are calculated is linked to this sheet.

Use the drop down menus to select the areas you want to model. You can choose the whole of Uganda, one or region, one or more district. If Uganda is selected then all districts are selected. If a region is selected all districts in the region are selected.

- The Uganda birth rate and infant mortality rate are combined to estimate number of deliveries.
- User can change per cent of deliveries that occur in public health facilities.
- Specifying whether a year is Baseline or SMGL will affect costing of medicines.
- Changing the number of deliveries per facility and preference of facility affects where in the health system cost of medicines are allocated.

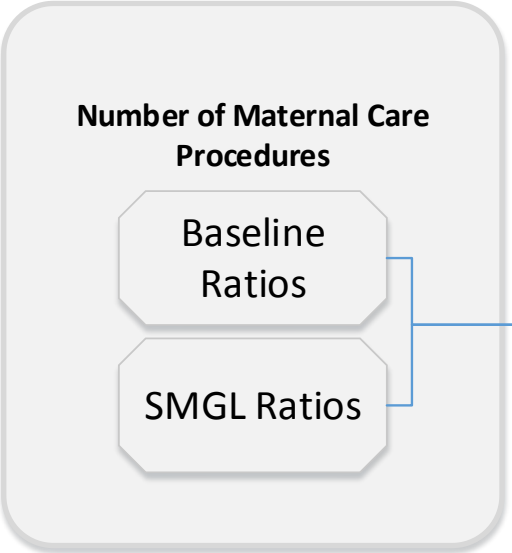

Costing of Ugandan Maternal Care
Ambu
Level III
Level IV
Hospitals
User can change number of ambulance drivers per Level 4 Facility and Hospitals. Assumptions about the number of ambulances can be modelled
Staffing assumptions per facility can be modelled at each facility. Basic assumptions about operating costs can be modelled. Important operating costs are set in the GenAssumptions sheet. Number of vehicles at facilities can be changed – the cost of operating the vehicles is set in GenAssumptions sheet.



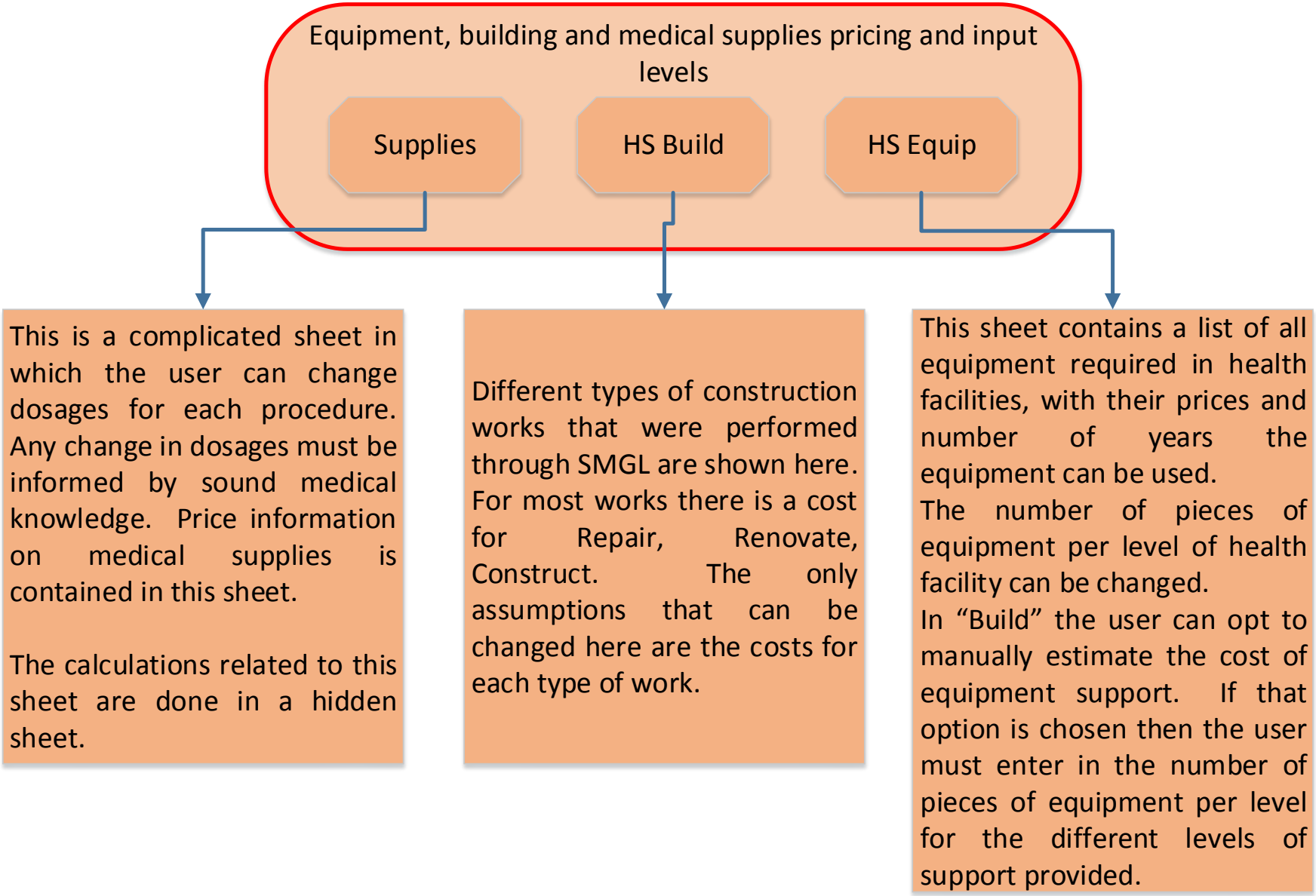
This contains data on populations and numbers of health facilities per district. There is no need to change anything in this sheet.

This sheet shows all the different types of health and programme workers used in the model. The salaries for each type of worker is shown here. In a lot of sheets the user can choose a type of worker from a drop down menu. The salary for all types of staff in the model are pulled from this sheet.

The data in this sheet should not be changed.



These sheets have the same structure. Baseline Ratios shows the ratios for years when SMGL has not been implemented.
The number of deliveries per level of the health system is determined by assumptions chosen in "Scenarios".
In this sheet the user can change the demand for each procedure (complication rate). The % chosen is relative to the number of deliveries.
The stocking rate of medicines can be set for each procedure. In the Comps Baseline sheet the stocking rate as a % of required medicines that government achieves is set. In SMGL Ratios the stocking rate that government with support from SMGL achieves is set.


Equipment, building and medical supplies pricing and input levels
Supplies
HS Build
HS Equip
This is a complicated sheet in which the user can change dosages for each procedure. Any change in dosages must be informed by sound medical knowledge. Price information on medical supplies is contained in this sheet.
The calculations related to this sheet are done in a hidden sheet.
Different types of construction works that were performed through SMGL are shown here. For most works there is a cost for Repair, Renovate, Construct. The only assumptions that can be changed here are the costs for each type of work.
This sheet contains a list of all equipment required in health facilities, with their prices and number of years the equipment can be used.
The number of pieces of equipment per level of health facility can be changed.
In "Build" the user can opt to manually estimate the cost of equipment support. If that option is chosen then the user must enter in the number of pieces of equipment per level for the different levels of support provided.

Costing of SMGL Activities (1 of 2)

SMGL Man

A range of simple assumptions about the number of lead companies, number of district offices for the lead company and staffing at head office and district level are set here.

Note that the number of lead companies is an absolute number will be set to zero if no districts are selected.

Aware

Assumptions about awareness creation campaigns are set in this sheet. Activities include demand creation workshops, drama groups, radio talk shows. There are a wide range of inputs the user can choose to include, or not, in the costing.

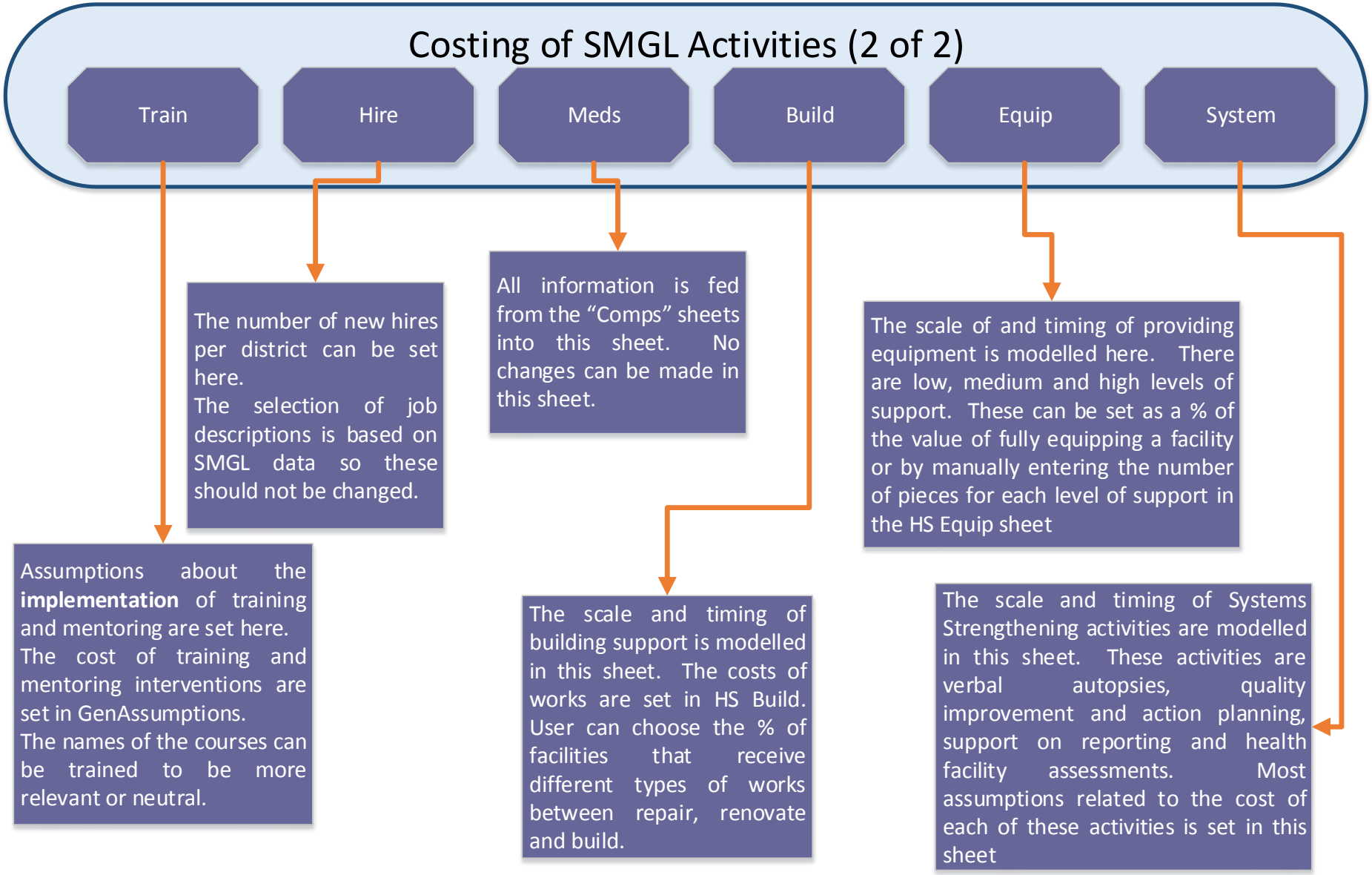
VHTs

Assumptions related to the number of Village Health Team per delivery, the structure of coordination and training of VHTs can be modelled.
Many options are provided so that different approaches to managing VHTs can be modelled.

TVs

Assumptions related to the value of TVs and the number purchased and redeemed and their value can be set. Number of TVs bought is driven by number each VHT buys.
Assumptions related to training of and incentives for Bodas can be modelled.

Ambu

Simple assumptions related to ambulance services are set here. Quantities are worked out per district.

Costing of SMGL Activities (2 of 2)