Technical Annex 1: Sampling

Sampling followed a four step procedure.

First, we selected the regions for inclusion in the analysis, which serve as the strata for further sampling. We selected regions for data collection based on variables we thought important for obtaining a representative sample (i.e. density of doctors, density of ART patient load per facility by region, and whether there would be facilities with doctors administering ART). We selected four regions of Ethiopia: Addis Ababa, Amhara, Benishangul Gumuz, and Oromiya.

Second, to ensure a comparison of different task shifting models for the analysis, we interviewed ART coordinators at the Federal Ministry of Health, Federal HIV/AIDS prevention and control office (FHPACO), and Regional Health Bureaus to identify hospitals where doctors were involved in delivering ART because these hospitals likely had the least amount of task shifting. We selected these networks with certainty, where a network consists of the hospital and the associated satellite health centers.

Third, from the list of remaining health networks within the selected regions, we selected healthcare networks using random sampling without replacement.

Fourth, we selected facilities within each sampled network. We selected the anchor hospital in each network with certainty together with a sample of 2 or 3 satellite health centers. We included only health centers that had been delivering ART as of the end of 2010 in the sampling framework to ensure two-year follow-up would be possible. If there were fewer than 4 eligible health centers in the network, then we selected all facilities. For networks with 4 or more eligible health centers, we further stratified health centers into those with a longer history of delivering ART services and those newer to ART service delivery. In the first instance, we randomly selected one clinic from the first stratum. Secondly, we randomly selected two additional clinics from the union of the second stratum and the balance from the first stratum. Substitution was used for nonresponse due to inaccessibility. This only occurred once. In this instance, the weight of originally selected clinic was simply given to its replacement.

Finally, within facilities where data on costs were collected, we also performed exit interviews with clients. We interviewed up to 15 ART clients per facility, but limited interviewing only to the one day. Thus, some facilities had fewer than 15 interviews.

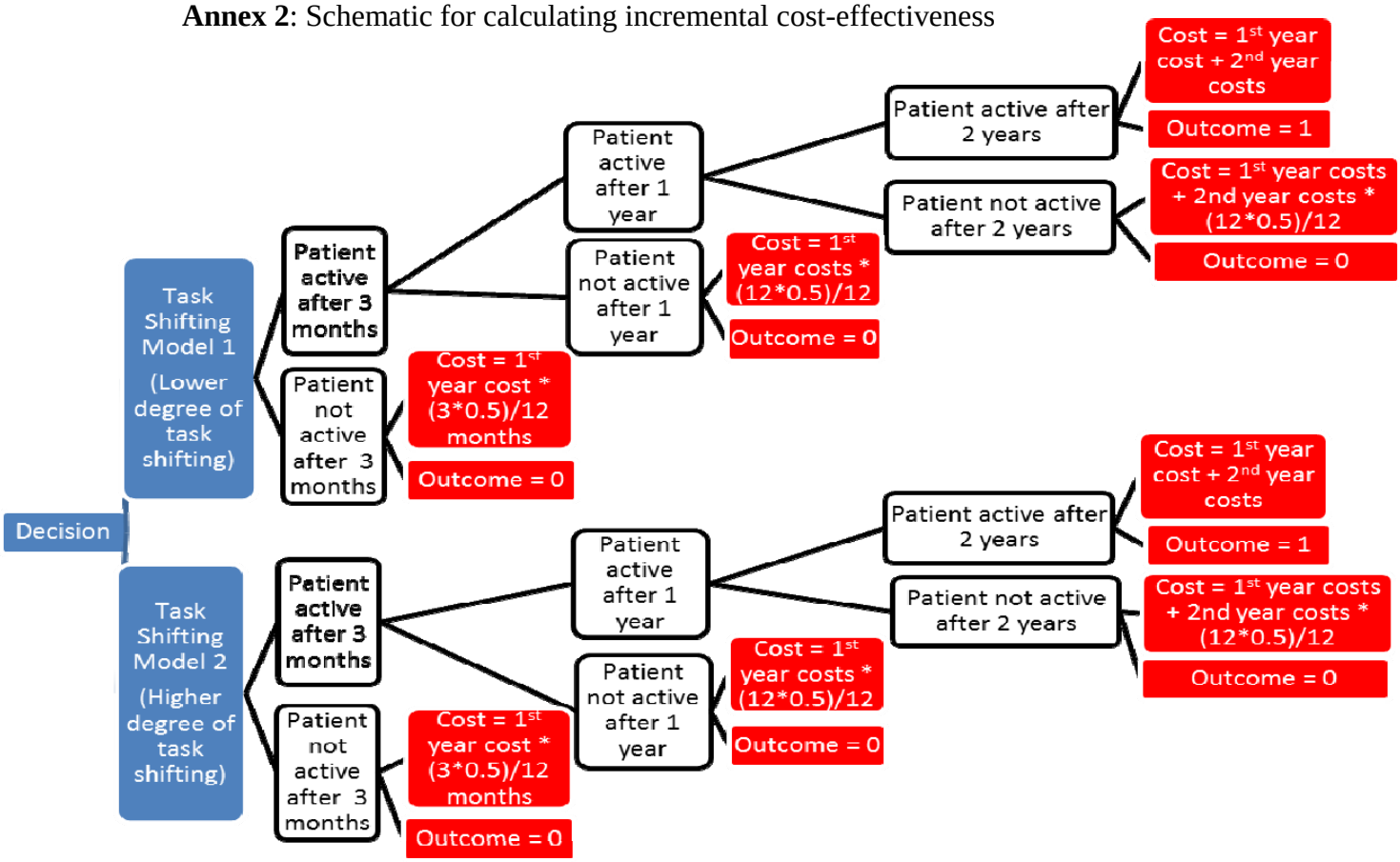
Using previous reports of treatment retention in Ethiopia,[30] we estimated the sample size assuming a design effect of 1.4, a difference in the probability of two year treatment retention of 0.07 between comparison arms, and a power of 0.8. However, due to the complicated nature of the sampling and analysis, we view these calculations as highly provisional. Table 1 shows the number of networks and facilities included in the analysis.

Annex Table S1. Network Sample Sizes

Region	Randomly Sampled Networks	Networks selected with certainty	Total studied networks	Visited Facilities
Addis Ababa	4 of 6	0	4	15
Amhara	4 of 12	5	9	32
Benishangul Gumuz*	0	2	2	8
Oromiya	4 of 28	2	6	24
Total	12 of 46	9	21	79

* One clinic was visited by local staff despite not being part of the sample or an official substitute. This "volunteer" clinic was given a weight of zero. Thus, the total sample size is 78 facilities, with 7 from Benishangul Gumuz.

Annex 2: Schematic for calculating incremental cost-effectiveness



Annex 3: Results of cost-effectiveness analysis

Comparison	Adjusted cost over 2 years per patient starting ART (risk adjusted)	Difference in risk-adjusted costs (95% CI*)	Proportion still active after 2 years	Difference in risk-adjusted effects (95% CI)	Cost-effectiveness ratio
Facilities with minimal or moderate task shifting (baseline)	$369	$36 (-40 to 111)	0.928	-0.004 (-0.009 to 0.002)	Not calculated: Maximal task shifting is more costly, less effective
Facilities with maximal task shifting	$404		0.926		

*Based on bootstrap standard error

Annex 4: Results of exploration of factors associated with outcome

Variable	**Reference group**	**Determinants of 2-year risk adjusted survival rate**				**Determinants of per patient activ years**	
		Model 1 with all variables		Model 2 without poorly measured variables		Model 3 with all variables	
		Beta	P < t	Beta	P < t	Beta	P < t
Maximal task shifting	Minimal or no task shifting	-0.01	0.70	-0.03	0.08	27.2	0.53
Facility reports support funding only coming from Ethiopian government	Multiple sources of funding	0.01	0.55			15.9	0.67
Facility where ART care is integrated with other kinds of care	Facility where ART is stand alone or delivered with other HIV care	-0.07	0.002**	-0.11	0.02*	-63.3	0.16
Facility started offering services after EFY 2000	Offered services in 1999 or earlier	-0.13	0.03*	-0.05	0.28	-34.6	0.22
Facility receiving referrals for some ART care	Facility does not	0.09	0.03*	0.10	0.09	-11.5	0.82

	receive referrals						
Facility referring patients up to [another] hospital for some ART care	Facility does not refer out	-0.12	0.001**	-0.14	0.002**	27.9	0.69
Facility reports offering ART mentoring to another facility	Facility does not report that it mentors	0.02	0.36			23.6	0.53
Average number of ART patients		0.0001	0.002**	0.0001	0.02*	-0.12	0.09
Square of the average number of ART patients		-8.91e-09	0.02*	-1.07e-08	0.06	0.00002	0.10
Facility located in urban area	Facility located in rural area	0.01	0.83	-0.02	0.69	-16.2	0.78
Facilities has a physician on staff	No physician	-0.10	0.006**	-0.13	0.007**	-81.9	0.13
Patients per clinical staff		0.00002	0.009**			0.01	0.25
Facility is a hospital	Health center	-0.07	0.009**	-0.03	0.34	83.1	0.01*
n		57		74		41	
R-squared		0.713		0.507		0.280	

The first two regressions use the risk adjusted probability of being active two years after ART initiation as the outcome variable. The first regression uses all available variables, while the second regression shows the results of a more parsimonious model using only variables with limited missing data. The third and fourth models repeat the regressions of the first two models, but use the 2-year cost per patient as the outcome variable.

* Significant at $p < 0.05$

** Significant at $p < 0.01$

Annex 5: Results of uncertainty analysis on incremental costs and effects after regression adjustment