Online Appendix

Surviving Bad News: Health Information Without Treatment Options

Alberto Ciancio[a,c*] Fabrice Kämpfen[b,c] Hans-Peter Kohler[c]
Rebecca Thornton[d]

[a]Adam Smith Business School, University of Glasgow
[b]School of Economics, University College Dublin
[c]Population Studies Center, University of Pennsylvania
[d]Hankamer School of Business, Baylor University

June 7, 2024

*E-mail address: alberto.ciancio@glasgow.ac.uk.

A Appendix

A.1 Instruments validity

The experiment randomized financial incentives and distance to the clinic, which we use as instrumental variables to predict whether an individual learns their HIV status. Ideally, the incentives and distance should be orthogonal to baseline individual characteristics. To test for orthogonality, we regress baseline individual characteristics of age, gender and HIV status, on the set of instruments. More specifically, we estimate the following specification: $X_i = Z_i\gamma + \alpha_d + \epsilon_i$ with $X_i = \{sex_i, age_i, hiv{+}_i\}$ and $Z_i = (any\_incentive_i, incentive_i, incentive_i^2, distance_i, distance_i^2)$. α_d are district fixed-effects and we cluster the standard errors at the village level. Results are shown in Table A.2.[1] None of the instruments predicts baseline characteristics of individuals in our sample, although we find a statistically significant correlation between distance and the probability of being HIV+. This is likely due in part to the smaller sample of HIV+ respondents and the way the randomization was implemented. The locations of the health care clinics depend on the geographical features of the villages and districts. The Northern region – Rumphi – is much more mountainous and sparse than the two other regions and has a lower prevalence of HIV+.[2]

We supplement our balance analysis by providing further evidence that our instruments are orthogonal to additional baseline characteristics that are not included in our main analysis. More specifically, Table A.5 provides more evidence of balance in baseline characteristics by showing the correlation between our set of instruments and marital status, whether respondents had any formal schooling, years of education, number of children, and a continuous wealth index based on a set of dwelling characteristics and ownership of household durable assets (constructed using first principal component analysis), all at baseline.

We show in Table A.6 that including these additional covariates in our main empirical specification barely changes the magnitude of the effect of learning HIV status on survival. We do lose some precision because we lose about 1,000 observations ($\sim$1/3 of our original sample)

The bottom panel of Table **??** shows the first stages of the IV regressions along with several tests that we perform to assess the strength of our instrumental variables. The detection of weak instruments in our setting is complicated by three factors: (i) we consider non-homoskedastic errors, (ii) we have several endogenous variables, (iii) we have more instrumental variables than endogenous variables (e.g., we are over-identified).

? recommends using the "Effective F-statistics" to detect weak instruments in the over-identified non-homoskedastic case with a single endogenous variable, developed by **?**. We have been unable to find equivalent statistics in the case with multiple endogenous variables (see footnote 4 of **?**) and

[1] We present reduced form estimates in Table A.3.

[2] To provide evidence that our main results are not biased due to the correlation between distance and HIV+ status of individuals, we show the results of a model that does not include distance as instrumental variables. We show in Table A.4 that our results are very similar to the main specification that also includes distance as instrumental variables. This suggests that our main results are not biased due to the correlation between distance and HIV+ status of individuals. We have nonetheless added this imbalance between distance and HIV status to our lists of caveats in the introduction.

no consensus, to our knowledge, has been reached on how to proceed in this context.[3] However, a recent working paper by **?** proposes a conservative testing procedure that extends the work by **?** and allows for more than one endogenous variables.

We follow **?** to assess the strength of our instrumental variables and report the following at the bottom of Table 2: (i) Cragg-Donald F-statistic: this statistic should be used in the case of multiple endogenous variables under homoscedastic errors. However, researchers usually report this statistic despite assuming non-homoscedastic errors when conducting inference for the second-stage parameters (**??**). (ii) Effective F-statistics by **?**, which are the preferred statistics in the over-identified case with non-homoscedastic errors and a single endogenous variable (**?**). To do so, we run separate first stage regressions to separately predict our two endogenous variables, receiving an HIV+ and HIV- diagnosis, using our set of instrumental variables. (iii) Lewis-Mertens statistics (g_{min}), which we compare to Imhof-based (**?**) critical values derived as in **?**.

As reported at the bottom of Table 2, Cragg-Donald and Effective F-statistics are well above the commonly-used rule-of-thumb value of 10 (**??**) and indicate that our instrumental variables are reasonably strong to predict our endogenous variables.

The bottom of Table 2 also reports cluster-robust g_{min}-statistics that accommodate multiple endogenous variables in an over-identified setting as defined in **?**. We compare these g_{min}-statistics to Imhof-based critical values defined at a significance level of $\alpha = 0.05$ and a relative bias threshold of $\tau = 0.2$ (**?**). The result of these conservative g_{min}-statistics-based tests suggests that in all but one case, we can reject the null hypothesis of weak instruments.[4]

Finally, we report Effective F-statistics for the specifications in which we estimate the effects of learning HIV status on survival separately for HIV+ and HIV– individuals. In this case, there is only one endogenous variable and the Effective F-statistics is the preferred statistics for detecting weak instruments (**?**). The Effective F-statistics reported at the bottom of each panel of Table A.7 are all above 10.[5]

A.2 Attrition in vital status

While the overall attrition rate in vital status, at 6% six years after the experiment, is notably low compared to studies in similar contexts, there is still a potential concern regarding differential attrition across treatment groups. However, our findings remain very robust to using econometric models that account for selective attrition. First, using our baseline instrumental variable specification, learning to be HIV– or HIV+ is unrelated to the availability of vital status information in any of the years these data are available (Table A.9). Second, our estimated treatment effects on

[3]**?** discusses inference with weak instruments and multiple endogenous regressors, in which case projection methods can be used to conduct tests and derive confidence sets, but this approach is known for suffering from poor power (**?**).

[4]More specifically, according to this test, our instruments slightly fall short of being strong enough (g_{min}=10.64 vs critical value of 11.18) in our 2006 specification. The instruments are however strong enough for the years when we estimate statistically significant effects of receiving HIV+ diagnosis on survival (2008, 2010, and 2018/2019).

[5]Table A.8 shows that excluding the interactions of Z with HIV+ (HIV–) when predicting "learning to be HIV— (HIV+)" leads to similar first stage estimates.

survival using our main specification outlined above are robust to using inverse probability weighting (Table A.10), where the weights are calculated using the probability of observing vital status information conditional on individuals controls and instruments for a given survey year.

Finally, another way to assess the importance of selective attrition is to model treatment effects under different assumptions about the unfound respondents. A commonly used approach is to report Manski extreme bounds, which impute minimum (maximum) outcome value for treated (control) respondents, and vice versa (**???**). We follow this approach by assigning vital status to respondents with missing data, and in particular by making extreme assumptions about the vital status of unfound HIV+ individuals. We assume the survival rate of unfound HIV− respondents to be on average the same as the one we observe in our data. This assumption is plausible given the very small magnitude in the effects of learning one's HIV− status on survival and attrition (Tables **??** and A.9), which suggests no selective attrition among HIV− respondents.

We therefore assign different survival rates for unfound respondents who are HIV−, HIV+ and learned about their status, and HIV+ and did not learn about their status. More specifically, we impute missing outcomes by setting the mortality rate of unfound HIV− respondents at 2.3% in 2008, 4.2% in 2010, and 11.2% in 2018 (see Table **??**) and estimate our treatment effects by considering *all the various possible combinations of mortality rates* among those who are HIV+ and learned about their status and those who are HIV+ and did not learn about their status, including the most extreme cases (i.e., all HIV+ who learned (did not learn) they were HIV+ died (were still alive) and vice versa). Practically, we simulate our data by randomly assigning living status to observations with missing outcomes a 100 times and report the average of the corresponding distributions of the treatment effects.

Figures A.2, A.3, and A.4 show the average treatment effects of our simulation exercise.[6] Each marker in these plots represents the average of 100 simulated treatment effects. The x-axis shows the survival rate among those who were HIV+ and did not learn about their HIV status while the figure legend reports the markers associated with the different mortality rates for HIV+ individuals who learn about their HIV status. The horizontal light (dark) grey line represents the effect of learning one's HIV− (HIV+) on survival estimated from our benchmark specification (Table **??**). One can see in Table A.2 that the effects of learning to be HIV+ on survival in 2008 are very consistent irrespective of the combinations of the survival rates among those who were HIV+ at baseline. The causal effects of learning one's HIV+ on survival monotonically increases (gets closer to 0) as we increase the survival rate among those who learned that they were HIV+ (while keeping constant the survival rate among those who did not learn their HIV+ status). The effects for those who learned that they were HIV- are relatively constant and remain close to our benchmark treatment effect (horizontal light grey line).

Results from the 2010 and 2018 simulations (Figures A.3 and A.4, respectively) show very similar

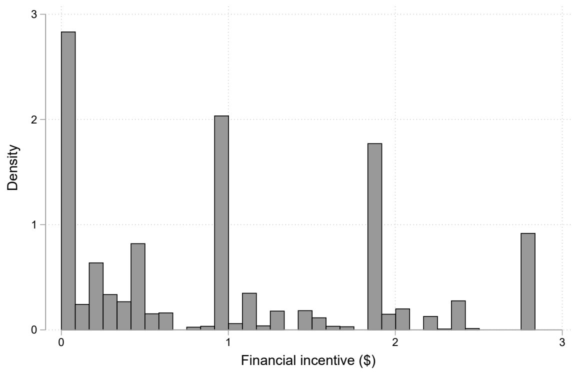
[6] For the 2008 simulation, among those with missing living status information, there were 8 individuals who were HIV+ and did not learn about their HIV status, and 5 who were HIV+ and learned about their HIV+ status. For the 2010 (2018) simulation, among those with missing living status information, there were 7 (12) individuals who were HIV+ and did not learn about their HIV status, and 5 (9) who were HIV+ and learned about their HIV+ status.

patterns than those presented in Figure A.2. This simulation exercise provides further evidence that selective attrition is unlikely to bias our treatment effects and reinforces the robustness of our benchmark results.

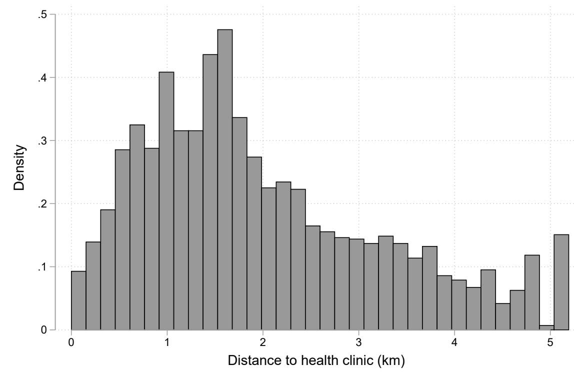
A.3 Figures

Figure A.1: Histograms of the instrumental variables used in this study

(a) Financial incentives

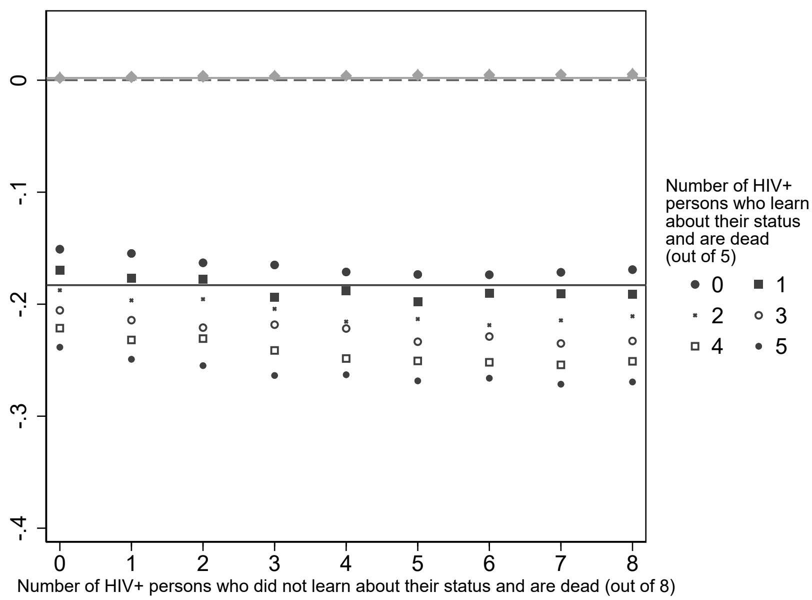


(b) Distance to health clinic



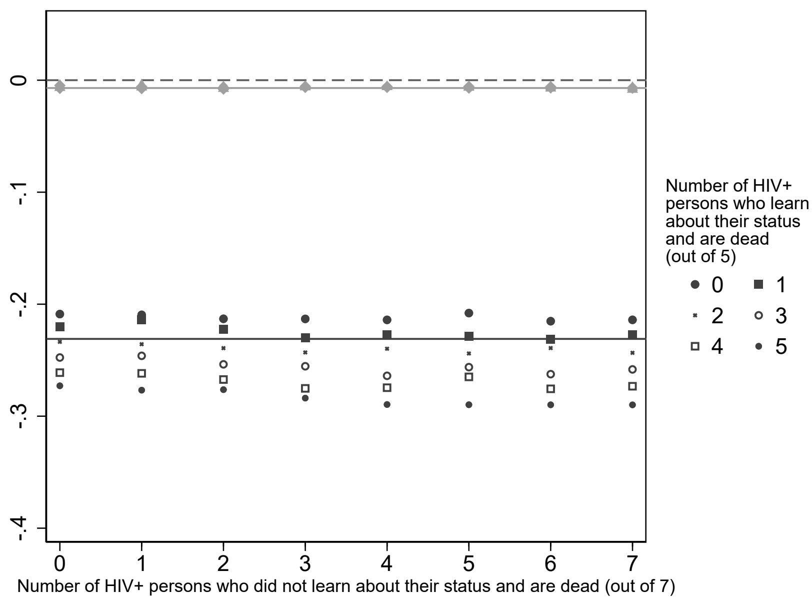
Note: The plot on the left (a) shows the histogram of the financial incentive in $ that respondents received in case they decided to receive their HIV test results. This plot of the right (b) shows the histogram of the distance between the place where the respondents live and the health clinic where respondents could pick up their HIV test results. The location of the health clinics was randomly selected in the respondents village where they live. We use both financial incentives and distance to health clinic as instrument variable to predict whether respondents learn about their HIV test results.

Figure A.2: **Causal effects estimated based on 100 simulations for each possible combination of mortality rates among unfound HIV+ individuals – 2008**



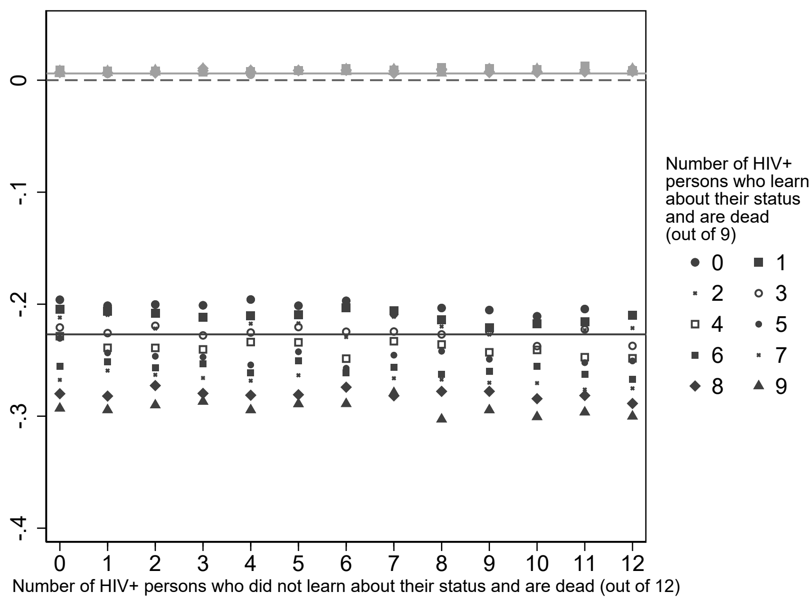
Note: The symbols represent the mean causal effects of learning HIV- (light color) and learning HIV+ (dark color) on the probability of being alive in 2008 based on our 100 simulations. There were 129 observations with missing living status in 2008, among which 116 were HIV- and 13 were HIV+. Among the 13 individuals who were HIV+, 5 learned about their status and 8 did not. We set the mortality rate of those who are HIV- at the rate we observe in our data (2.3% between 2004 and 2008). Symbols (in the legend) represent different mortality rate among individuals who were HIV+ and learned about their HIV status. The x-axis shows the mortality rate among those who were HIV+ and did not learn about their HIV status. The horizontal lines represent our benchmark estimates reported in Table ??.

Figure A.3: **Causal effects estimated based on 100 simulations for each possible combination of mortality rates among unfound HIV+ individuals – 2010**



Note: The symbols represent the mean causal effects of learning HIV- (light color) and learning HIV+ (dark color) on the probability of being alive in 2010 based on our 100 simulations. There were 174 observations with missing living status in 2010, among which 162 were HIV- and 12 were HIV+. Among the 12 individuals who were HIV+, 7 learned about their status and 5 did not. We set the mortality rate of those who are HIV- at the rate we observe in our data (4.2% between 2004 and 2010). Symbols (in the legend) represent different mortality rate among individuals who were HIV+ and learned about their HIV status. The x-axis shows the mortality rate among those who were HIV+ and did not learn about their HIV status. The horizontal lines represent our benchmark estimates reported in Table ??.

Figure A.4: **Causal effects estimated based on 100 simulations for each possible combination of mortality rates among unfound HIV+ individuals – 2018-2019**



Note: The symbols represent the mean causal effects of learning HIV- (light color) and learning HIV+ (dark color) on the probability of being alive in 2018 based on our 100 simulations. There were 491 observations with missing living status in 2018-2019, among which 470 were HIV- and 21 were HIV+. Among the 21 individuals who were HIV+, 9 learned about their status and 12 did not. We set the mortality rate of those who are HIV- at the rate we observe in our data (11.2% between 2004 and 2018). Symbols (in the legend) represent different mortality rate among individuals who were HIV+ and learned about their HIV status. The x-axis shows the mortality rate among those who were HIV+ and did not learn about their HIV status. The horizontal lines represent our benchmark estimates reported in Table ??.

A.4 Tables

- Table A.1: Summary statistics of the 2005 study sample
- Table A.2: Exogeneity of the instruments
- Table A.3: Reduced form estimates
- Table A.4: The effects of learning HIV status on survival using only financial incentives as IVs
- Table A.5: Exogeneity of instruments with other characteristics at baseline
- Table A.6: The effects of learning HIV status on survival including an extended set of control characteristics at baseline
- Table A.7: The effects of learning HIV status on survival by HIV status
- Table A.8: First stage regressions by survey year excluding the interactions of Z with HIV+ (HIV–) when predicting "learning to be HIV— (HIV+)"
- Table A.9: The effects of learning HIV status on the availability of vital status information
- Table A.10: Attrition corrected treatment effects: inverse probability weights
- Table A.11: The effects of learning HIV status on economic worries
- Table A.12: The effects of learning HIV status on economic uncertainties
- Table A.13: The effects of learning HIV status on drinking behaviors

Table A.1: Summary statistics of the 2005 study sample

	All			***HIV−***			***HIV+***		
	Mean	Sd	N	Mean	Sd	N	Mean	Sd	N
Sum of worries	2.423	1.028	1474	2.401	1.017	1408	2.879	1.144	66
Worries about AIDS-related health problem	0.071	0.256	1471	0.063	0.242	1405	0.242	0.432	66
Worries about other health problems	0.372	0.484	1472	0.364	0.481	1406	0.545	0.502	66
Worries about paying school fees	0.209	0.407	1470	0.205	0.404	1404	0.288	0.456	66
Worries about finding or earning enough money	0.931	0.254	1472	0.930	0.255	1406	0.939	0.240	66
Worries about having enough food	0.828	0.378	1465	0.828	0.378	1399	0.833	0.376	66
Sum economic uncertainties	4.513	1.117	1474	4.513	1.118	1408	4.500	1.099	66
Not easy to get MKW500	0.864	0.343	1474	0.864	0.343	1408	0.879	0.329	66
Future is uncertain	0.747	0.435	1473	0.746	0.436	1407	0.773	0.422	66
Family not enough food	0.693	0.461	1464	0.693	0.461	1399	0.692	0.465	65
Not enough for basic essential	0.910	0.287	1472	0.910	0.287	1406	0.909	0.290	66
Only God determines future	0.967	0.180	1471	0.967	0.180	1405	0.970	0.173	66
Children take care of my health when old	0.341	0.474	1472	0.344	0.475	1406	0.288	0.456	66
Time preference - USD in a month (vs 4.73 USD today)	10.855	10.286	1427	10.923	10.367	1364	9.366	8.269	63
Drink past 12 months	0.128	0.334	1441	0.125	0.331	1375	0.182	0.389	66
# days drunk last week	0.165	0.746	1411	0.166	0.754	1347	0.141	0.560	64
At least 1 day drunk last week	0.063	0.243	1411	0.062	0.241	1347	0.094	0.294	64
At least 2 days drunk last week	0.043	0.202	1411	0.044	0.205	1347	0.016	0.125	64
Drunk during weekdays (except Friday)	0.052	0.222	1411	0.051	0.221	1347	0.062	0.244	64
Subjective life expectancy at age	26.383	16.006	1270	26.668	16.065	1218	19.712	12.985	52

Note: The sample is derived from all respondents who got tested for HIV in 2004 and were interviewed in 2005 as part of the intervention follow-up survey. Information was collected only among respondents living in Rumphi and Balaka. "Sd" stands for standard errors and "N" for the number of observations. "Worries" and "uncertainties" are all dichotomous variables. Sum of "worries" also includes "Other" worries that respondents could report. 'Drunk on weekdays" is a dichotomous variable that takes the value 1 if a respondent admitted to have been drunk on Monday, Tuesday, Wednesday, and/or Thursday. Variables in **bold** are part of the main mechanisms reported in Table **??**.

Table A.2: Exogeneity of the instruments

	Male (1)	Age (2)	HIV+ (3)
Any incentive	-0.050	1.342	-0.003
	(0.039)	(1.032)	(0.018)
Incentive	0.062	0.379	-0.013
	(0.057)	(1.381)	(0.022)
Incentive squared	-0.017	-0.315	0.004
	(0.019)	(0.451)	(0.008)
Distance in km	-0.028	-0.234	-0.032
	(0.020)	(0.971)	(0.016)
Distance squared	0.007	0.118	0.005
	(0.004)	(0.195)	(0.003)
Constant	0.482	37.442	0.113
	(0.027)	(1.154)	(0.023)
Observations	2823	2823	2823
R^2	0.002	0.005	0.006

Note: The table shows estimates of the effect on the instruments on predetermined variables. Reported in parentheses are standard errors that are clustered at the village level. All specifications control for region fixed effects.

Table A.3: Reduced form estimates

Respondent is alive in:	2006	2008	2010	2018-19
	(1)	(2)	(3)	(4)
Any incentive x HIV–	0.016	0.016	0.020	0.014
	(0.005)	(0.012)	(0.015)	(0.026)
Incentive x HIV–	-0.019	-0.005	-0.014	-0.019
	(0.008)	(0.015)	(0.017)	(0.030)
Incentive squared x HIV–	0.005	-0.000	0.002	0.007
	(0.003)	(0.005)	(0.006)	(0.010)
Distance in km x HIV–	0.006	0.009	0.008	0.004
	(0.006)	(0.008)	(0.010)	(0.018)
Distance squared x HIV–	-0.002	-0.003	-0.002	-0.001
	(0.001)	(0.002)	(0.002)	(0.003)
Any incentive x HIV+	-0.017	-0.095	-0.124	-0.244
	(0.083)	(0.105)	(0.112)	(0.119)
Incentive x HIV+	-0.072	-0.068	-0.098	0.099
	(0.130)	(0.177)	(0.175)	(0.168)
Incentive squared x HIV+	0.023	0.036	0.048	-0.042
	(0.043)	(0.054)	(0.055)	(0.057)
Distance in km x HIV+	-0.048	-0.097	-0.154	-0.044
	(0.056)	(0.085)	(0.093)	(0.122)
Distance squared x HIV+	0.011	0.021	0.032	0.008
	(0.011)	(0.017)	(0.018)	(0.021)
Observations	2756	2694	2649	2332

Note: The table shows estimates of the effects of the set of instrumental variables on survival. Reported in parentheses are standard errors that are clustered at the village level. All specifications control for HIV status, sex, age and its square, and region fixed effects.

Table A.4: The effects of learning HIV status on survival using only financial incentives as IVs

Respondent is alive in:	***Second stage estimates***			
	2006	2008	2010	2018-19
	IV	IV	IV	IV
	(1)	(2)	(3)	(4)
Learn HIV–	0.002	0.003	-0.003	0.006
	(0.007)	(0.014)	(0.015)	(0.024)
Learn HIV+	-0.166	-0.217	-0.293	-0.282
	(0.094)	(0.112)	(0.114)	(0.144)
Observations	2756	2694	2649	2332

	First stage estimates							
	Learn HIV–	Learn HIV+	Learn HIV–	Learn HIV+	Learn HIV–	Learn HIV+	Learn HIV–	Learn HIV+
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Any incentive x HIV–	0.274	-0.001	0.274	-0.002	0.280	-0.002	0.270	-0.003
	(0.030)	(0.002)	(0.031)	(0.002)	(0.032)	(0.002)	(0.034)	(0.002)
Incentive x HIV–	0.250	0.001	0.257	0.001	0.256	0.001	0.265	0.002
	(0.038)	(0.001)	(0.040)	(0.001)	(0.040)	(0.001)	(0.042)	(0.001)
Incentive squared x HIV–	-0.055	-0.000	-0.056	-0.000	-0.057	-0.000	-0.059	-0.000
	(0.012)	(0.000)	(0.012)	(0.000)	(0.012)	(0.000)	(0.013)	(0.000)
Any incentive x HIV+	-0.023	0.201	-0.021	0.186	-0.025	0.196	-0.024	0.184
	(0.016)	(0.122)	(0.016)	(0.126)	(0.016)	(0.122)	(0.016)	(0.136)
Incentive x HIV+	0.023	0.480	0.026	0.546	0.030	0.537	0.028	0.558
	(0.018)	(0.142)	(0.019)	(0.152)	(0.019)	(0.148)	(0.019)	(0.155)
Incentive squared x HIV+	-0.005	-0.157	-0.006	-0.177	-0.007	-0.179	-0.007	-0.179
	(0.006)	(0.047)	(0.006)	(0.049)	(0.006)	(0.048)	(0.006)	(0.050)
Observations	2756	2756	2694	2694	2649	2649	2332	2332
Effective F-statistics	244.71	14.89	233.52	16.29	246.38	15.86	220.54	15.07
Cragg-Donald F-statistics	117.29		123.29		124.25		109.58	
Lewis-Mertens statistics (g_{min})	15.01		16.42		15.98		15.19	
g_{min} critical values ($\alpha = 0.05, \tau = 0.2$)	12.51		12.42		12.35		12.40	

Note: The first panel shows the estimates of the effect of learning HIV status on the probability of being alive for various years. The second panel shows the first stage estimates. All these estimations are derived using only financial incentives as instrumental variables. Reported in parentheses are standard errors that are clustered at the village level. All specifications control for HIV status, sex, age and its square, and region fixed effects. Effective F-statistics are based on **?**. Lewis-Mertens statistics (g_{min}) are based on **?** and robust to clustering. g_{min} critical values are set at a 5% significance level ($\alpha = 0.05$) and 20% relative bias threshold ($\tau = 0.2$) (**?**). These critical values are derived from clustered errors and using an Imhof distribution (**?**).

Table A.5: Exogeneity of instruments with other characteristics at baseline

	Characteristics of individuals at baseline							
	Married	Any schooling	Any schooling (extended)	Years of schooling	Years of schooling (extended)	Number of children	Wealth index	Wealth index quantile
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Any incentive	-0.028	-0.045	-0.054	-0.188	-0.228	0.283	-0.100	-0.107
	(0.031)	(0.038)	(0.034)	(0.281)	(0.275)	(0.267)	(0.142)	(0.093)
Incentive	0.005	0.006	0.011	0.018	0.041	0.005	0.078	0.071
	(0.042)	(0.049)	(0.046)	(0.383)	(0.377)	(0.312)	(0.194)	(0.126)
Incentive squared	-0.005	0.006	0.004	0.003	0.001	-0.028	-0.027	-0.023
	(0.014)	(0.015)	(0.015)	(0.116)	(0.114)	(0.103)	(0.068)	(0.043)
Distance in km	-0.008	-0.005	0.007	0.292	0.280	-0.013	0.116	0.026
	(0.036)	(0.040)	(0.038)	(0.291)	(0.282)	(0.195)	(0.170)	(0.100)
Distance squared	0.003	0.002	0.001	-0.052	-0.049	0.009	-0.034	-0.010
	(0.006)	(0.009)	(0.008)	(0.060)	(0.058)	(0.035)	(0.032)	(0.020)
Constant	0.784	0.818	0.814	3.529	3.632	3.864	-0.443	
	(0.041)	(0.044)	(0.040)	(0.307)	(0.291)	(0.284)	(0.200)	
Observations	2693	2538	2778	2026	2068	1999	1952	1952
R^2	0.004	0.137	0.128	0.164	0.159	0.014	0.167	

Note: The table shows estimates of the effect on the instruments on predetermined variables. Reported in parentheses are standard errors that are clustered at the village level. All specifications control for region fixed effects. The extended measures of schooling include information from the 2001 and 1998 waves. Other measures only include information from the 2004 wave. "Wealth index" is a continuous wealth index based on a set of dwelling characteristics and ownership of household durable assets, constructed using first principal component analysis (**???**). The last column report estimates from ordered probit regression instead of linear regression.

Table A.6: The effects of learning HIV status on survival including an extended set of control characteristics at baseline

Respondent is alive in:	2006	2008	2010	2018-19
	(1)	(2)	(3)	(4)
Learn HIV–	0.005	0.013	-0.003	0.012
	(0.009)	(0.022)	(0.022)	(0.031)
Learn HIV+	-0.064	-0.190	-0.225	-0.258
	(0.108)	(0.148)	(0.141)	(0.136)
Married	-0.014	0.020	0.031	0.081
	(0.010)	(0.021)	(0.024)	(0.034)
Number of children	0.002	0.003	0.007	0.008
	(0.001)	(0.002)	(0.002)	(0.004)
Any schooling (extended)	0.002	-0.015	-0.024	-0.004
	(0.006)	(0.011)	(0.016)	(0.018)
Wealth index	0.001	0.001	-0.000	-0.000
	(0.001)	(0.002)	(0.002)	(0.003)
Observations	1774	1753	1729	1630
R^2	0.161	0.111	0.100	0.149

Note: The table shows estimates of the effect on the probability of being alive for various years. Reported in parentheses are standard errors that are clustered at the village level. All specifications control for HIV status, sex, age and its square, region fixed effects, as well as whether the respondent is married at baseline, whether the respondent has any formal schooling (extended measure including information from the 2001 and 1998 waves), the number of children at baseline, and a continuous wealth index based on a set of dwelling characteristics and ownership of household durable assets, constructed using first principal component analysis (**???**). The instruments to predict whether respondents learned their HIV status are: whether respondents received any financial incentive, the amount of the incentive, the square of the amount of the incentive, the distance to the health clinic where respondents can obtain their results, the square distance, as well as all of their interaction with HIV status.

Table A.7: The effects of learning HIV status on survival by HIV status

Respondent is alive in:	2006	2008	2010	2018-19
	(1)	(2)	(3)	(4)
HIV–				
Learn HIV–	-0.000	-0.003	-0.011	0.005
	(0.007)	(0.014)	(0.014)	(0.024)
Observations	2589	2534	2488	2180
R^2	0.056	0.049	0.039	0.086
Effective F-statistics	131.55	131.56	133.80	109.91
HIV+				
Learn HIV+	-0.084	-0.115	-0.164	-0.188
	(0.069)	(0.105)	(0.107)	(0.137)
Observations	167	160	161	152
R^2	0.429	0.225	0.187	0.118
Effective F-statistics	10.34	11.33	11.93	12.84

Note: The table shows estimates of the effect on the probability of being alive for various years. Reported in parentheses are standard errors that are clustered at the village level. All specifications control for sex, age and its square, and region fixed effects. The instruments to predict whether respondents learned their HIV status are: whether respondents received any financial incentive, the amount of the incentive, the square of the amount of the incentive, the distance to the health clinic where respondents can obtain their results and the square distance. The first (second) panel restricts the analysis to individuals who are HIV– (HIV+) at baseline. Effective F-statistics are based on **?**.

Table A.8: First stage regressions by survey year excluding the interactions of Z with HIV+ (HIV–) when predicting "learning to be HIV— (HIV+)"

	2006		***2008***		***2010***		***2018-19***	
	Learn HIV–	Learn HIV+	Learn HIV–	Learn HIV+	Learn HIV–	Learn HIV+	Learn HIV–	Learn HIV+
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Any incentive x HIV–	0.275		0.274		0.280		0.271	
	(0.031)		(0.032)		(0.032)		(0.034)	
incentive x HIV–	0.252		0.258		0.258		0.265	
	(0.038)		(0.041)		(0.040)		(0.042)	
Incentive squared x HIV–	-0.055		-0.057		-0.057		-0.059	
	(0.012)		(0.012)		(0.012)		(0.013)	
Distance in km x HIV–	-0.046		-0.047		-0.043		-0.047	
	(0.029)		(0.029)		(0.030)		(0.034)	
Distance squared x HIV–	0.006		0.006		0.005		0.006	
	(0.005)		(0.005)		(0.006)		(0.006)	
Any incentive x HIV+		0.215		0.211		0.222		0.217
		(0.118)		(0.119)		(0.114)		(0.126)
Incentive x HIV+		0.468		0.518		0.509		0.511
		(0.140)		(0.149)		(0.144)		(0.153)
Incentive squared x HIV+		-0.149		-0.165		-0.167		-0.160
		(0.046)		(0.047)		(0.046)		(0.049)
Distance in km x HIV+		-0.235		-0.250		-0.277		-0.318
		(0.104)		(0.095)		(0.091)		(0.085)
Distance squared x HIV+		0.041		0.044		0.049		0.058
		(0.019)		(0.018)		(0.017)		(0.016)
Constant	0.326	-0.017	0.328	-0.014	0.314	-0.012	0.292	-0.012
	(0.068)	(0.018)	(0.070)	(0.018)	(0.070)	(0.019)	(0.079)	(0.020)
Observations	2756	2756	2694	2694	2649	2649	2332	2332

Note: The table shows estimates of the first stage. These estimates are derived from models that exclude the interactions of Z with HIV+ (HIV–) when predicting "learning to be HIV— (HIV+)". Reported in parentheses are standard errors that are clustered at the village level. All specifications control for sex, age and its square, and region fixed effects.

Table A.9: The effects of learning HIV status on the availability of vital status information

Living status information available in:				
	2006	2008	2010	2018-19
	(1)	(2)	(3)	(4)
Learn HIV−	0.003	-0.013	-0.023	0.011
	(0.010)	(0.016)	(0.016)	(0.031)
Learn HIV+	0.010	-0.074	-0.074	0.040
	(0.038)	(0.079)	(0.076)	(0.094)
Observations	2823	2823	2823	2823

Note: The table shows estimates of the effect on the probability of having vital status information for various years. Reported in parentheses are standard errors that are clustered at the village level. All specifications control for HIV status, sex, age and its square, and region fixed effects. The instruments to predict whether respondents learned their HIV status are: whether respondents received any financial incentive, the amount of the incentive, the square of the amount of the incentive, the distance to the health clinic where respondents can obtain their results, the square distance, as well as all of their interaction with HIV status.

Table A.10: Attrition corrected treatment effects: Inverse probability weights

Respondent is alive in:				
	2006	2008	2010	2018-19
	(1)	(2)	(3)	(4)
Learn HIV−	0.002	0.002	-0.008	0.008
	(0.006)	(0.014)	(0.014)	(0.024)
Learn HIV+	-0.112	-0.182	-0.226	-0.238
	(0.089)	(0.110)	(0.112)	(0.128)
Observations	2756	2694	2649	2332

Note: The table shows estimates of the effect on the probability of being alive for various years when applying inverse probability weights. Reported in parentheses are standard errors that are clustered at the village level. All specifications control for HIV status, sex, age and its square, and region fixed effects. The instruments to predict whether respondents learned their HIV status are: whether respondents received any financial incentive, the amount of the incentive, the square of the amount of the incentive, the distance to the health clinic where respondents can obtain their results, the square distance, as well as all of their interaction with HIV status.

Table A.11: The effects of learning HIV status on economic worries

Worried about:	AIDS-related health problems (1)	Other health problems (2)	Paying school fees (3)	Having enough food (4)	Having enough money (5)	**Sum of worries** (6)
Learn HIV−	-0.083	-0.121	0.067	-0.066	-0.016	-0.289
	(0.030)	(0.064)	(0.046)	(0.046)	(0.025)	(0.133)
Learn HIV+	0.297	0.396	-0.129	0.266	-0.013	0.840
	(0.083)	(0.172)	(0.142)	(0.157)	(0.026)	(0.363)
Observations	1471	1472	1470	1465	1472	1474

Note: The sum of worries includes "Other" worries that respondents could report. The table shows estimates of the effect on various worries that respondents reported in 2005. We control for the number of components that are included in the sum. Reported in parentheses are standard errors that are clustered at the village level. All outcomes variables were measured in 2005 as part of a follow-up survey a few months after the intervention. Details about the outcome variables used including sample descriptive statistics are detailed in Appendix Table A.4. All specifications control for HIV status, sex, age and its square, and region fixed effects. The instruments to predict whether respondents learned their HIV status are: whether respondents received any financial incentive, the amount of the incentive, the square of the amount of the incentive, the distance to the health clinic where respondents can obtain their results, the square distance, as well as all of their interaction with HIV status.

Table A.12: The effects of learning HIV status on economic uncertainties

Economic uncertainties:	Not easy to get MKW500 (1)	Future uncertain (2)	Not enough food (3)	Not enough for essentials (4)	God determine future (5)	Children take care (6)	**Sum of economic uncertainties** (7)
Learn HIV−	0.001	-0.085	-0.166	-0.034	0.026	0.017	-0.185
	(0.031)	(0.042)	(0.048)	(0.035)	(0.018)	(0.063)	(0.243)
Learn HIV+	0.141	0.081	0.387	0.009	0.072	0.134	0.934
	(0.110)	(0.154)	(0.164)	(0.075)	(0.031)	(0.215)	(0.390)
Observations	1474	1473	1464	1472	1471	1472	1474

Note: "Children" includes both "sons"' and "daughters"' help. The table shows estimates of the effect on various economic uncertainties that respondents reported in 2005. We control for the number of components that are included in the sum. Reported in parentheses are standard errors that are clustered at the village level. All outcomes variables were measured in 2005 as part of a follow-up survey a few months after the intervention. Details about the outcome variables used including sample descriptive statistics are detailed in Appendix Table A.4. All specifications control for HIV status, sex, age and its square, and region fixed effects. The instruments to predict whether respondents learned their HIV status are: whether respondents received any financial incentive, the amount of the incentive, the square of the amount of the incentive, the distance to the health clinic where respondents can obtain their results, the square distance, as well as all of their interaction with HIV status.

Table A.13: The effects of learning HIV status on drinking behaviors

	Drink alcohol past 12 months	# of days drunk last week	At least 1 day drunk last week	At least 2 days drunk last week	Drunk on weekdays
Learn HIV–	0.057	0.074	0.035	0.026	-0.007
	(0.038)	(0.074)	(0.028)	(0.018)	(0.024)
Learn HIV+	0.211	0.188	0.045	0.046	0.099
	(0.102)	(0.126)	(0.070)	(0.029)	(0.041)
Observations	1441	1411	1411	1411	1411

Note: "Drunk on weekdays" is a dichotomous variable that takes the value 1 if a respondent admitted to have been drunk on Monday, Tuesday, Wednesday, and/or Thursday. The table shows estimates of the effect on drinking behaviors. Reported in parentheses are standard errors that are clustered at the village level. All outcomes variables were measured in 2005 as part of a follow-up survey a few months after the intervention. Details about the outcome variables used including sample descriptive statistics are detailed in Appendix Table A.4. All specifications control for HIV status, sex, age and its square, and region fixed effects. The instruments to predict whether respondents learned their HIV status are: whether respondents received any financial incentive, the amount of the incentive, the square of the amount of the incentive, the distance to the health clinic where respondents can obtain their results, the square distance, as well as all of their interaction with HIV status.