

Online Appendix

A Additional information on experimental procedures

A.1 Further intervention details

The STYL manual is available online, including day-by-day breakdowns of the curriculum and homework. See <http://chrisblattman.com/documents/policy/2015.STYL.Program.Manual.pdf>

There were virtually no formally-trained psychologists or counselors in the country of Liberia at the time, and so all sessions were led by facilitators trained by NEPI. Generally speaking, they had limited formal education or counseling experience. Their experiences, however, made them natural role models for the students, as they modelled the desired changes in behavior and identity.

The curriculum focused on three main forms of change:

Changing social identity In the early weeks, facilitators encouraged men to try to maintain some simple behaviors associated with their new social identity. This included getting a haircut, wearing shoes and pants instead of sandals and shorts, improving personal hygiene and the cleanliness of their living area, and reducing substance abuse. As the men progressed, facilitators also encouraged men to engage with society in planned and unaccustomed ways, akin to exposure therapy. For instance, homework included reintroducing themselves to their family, joining community sports, and visiting banks, supermarkets, shops, and so forth. Men also studied successful people in their community and reached out to one as a mentor. Men then discussed their successes and failures as a group.

Future orientation In the middle and final weeks, the men practiced breaking down large goals into smaller ones, and then creating plans to accomplish them via concrete steps. For homework they would attempt some of these plans. Examples include how to feed their family the next day, starting a vegetable garden, making a savings plan, reconciling with estranged family, or starting a business. They began with easier assignments and increased in difficulty with time.

NEPI offered no incentives to attend except for lunch, and the subjects were unaware of the cash grants program until after the therapy was complete.

Self control Throughout the nine weeks, men worked on emotional regulation and anger management. During the group sessions, the men discussed angry and hostile thoughts,

emotions, and reactions. They learned to recognize and connect them to bad outcomes, like violence and exclusion. They also began to practice techniques to manage these automatic responses. For instance, they practiced social skills for managing threatening situations, and learned techniques to calm oneself, such as walking away, doing other activities, or breathing techniques.

A.2 Target population and recruitment

We recruited subjects in five mixed-income residential neighborhoods in Monrovia with large populations of high-risk men. Within these neighborhoods, NEPI staff focused on areas known for criminal and antisocial behavior, including known drug selling areas, areas for the fencing and resale of stolen goods, informal settlements with high levels of interpersonal street violence, and so forth. Recruiters also approached men that were homeless, drug-using, or appeared disreputable in appearance.

When they approached a new subject, recruiters described the purpose of the program (transition to a less violent, noncriminal lifestyle), the therapy, the allocation by lottery, and the associated research, and asked subjects if they wanted to participate. Recruiters never mentioned cash grants.

About one third of those approached declined. Some refused because they felt they were the wrong target—poor but not engaged in criminal or violent activities. Others were mistrustful or were content with their current illicit activities. We were unable to collect survey data on or track these individuals that refused any engagement.

Note that these procedures tried to minimize the possibility of spillovers between treatment and control subjects. To do so, we designed recruitment to be highly dispersed. Each of the five neighborhoods had a population of roughly 100,000, including at least a thousand high-risk young men. To maximize independence of the study subjects (and reduce potential for spillovers) recruiters approached just one out of every 7–10 high-risk men they visually identified. Over several weeks, recruiters identified roughly 8,000 potentially high-risk men and approached 1,500. In our five neighborhoods, we estimate that our sample of 999 men represents 0.6% of all adult males and about 12% of all men aged 18–35 who are in the bottom decile of income.

When a subject agreed to participate in the study, NEPI immediately introduced them to a representative of a nonprofit research organization, Innovations for Poverty Action (IPA), for a baseline survey and random assignment to the first intervention, therapy.

In the end, the study recruited 999 men. On average they were 25 years of age, had nearly eight years of schooling, earned about \$68 in the month previous to the baseline

Table A.1: Study sample and treatment assignment by randomization block

	Round 1					Round 7				
	#	Therapy (%)	Cash (%)	Therapy cash (%)	Control (%)	#	Therapy (%)	Cash (%)	Therapy cash (%)	Control (%)
Red light	100	28.0	24.0	25.0	23.0	72	25.0	23.6	30.6	20.8
Red light, second phase	219	26.9	25.1	24.2	23.7	182	26.4	24.2	25.8	23.6
Central Monrovia	179	31.8	19.0	31.8	17.3	157	29.3	19.7	33.1	17.8
Clara Town	175	28.6	27.4	22.9	21.1	140	26.4	28.6	22.1	22.9
Logan Town	86	26.7	29.1	19.8	24.4	67	23.9	28.4	22.4	25.4
New Kru Town	240	26.3	26.7	23.8	23.3	215	26.0	27.0	25.1	21.9

survey, worked an average of 46 hours per week (mainly in low skill labor and illicit work), and had \$34 informally saved. Only 38% were former members of armed militant groups.

A.3 Randomization and balance

Initially, recruited subjects were randomly assigned to therapy via a draw, where they drew colored chips from an opaque bag (with replacement). We did so partly for transparency and trust, and partly to eliminate the need to recontact these hard-to-find individuals post-randomization. Therapy began the week following randomization.

About one week after the final week of therapy, NEPI re-contacted all subjects and told them that another opportunity was available to them: a lottery for cash grants. The draws were conducted by an international nonprofit named Global Communities, in cooperation with NEPI. Again, one by one, subjects selected colored chips from a bag. For safety, they drew their lot and received the funds in private, and were immediately transported to a location of their choice by motorbike taxi. For those who did not attend, a chip was drawn on their behalf, and NEPI attempted to track them down. Of those assigned to a grant, 98% received it.

For logistical reasons we conducted this experiment in three phases of 100, 398, and 501 subjects between December 2010 and March 2012. Over the 3 phases, the experiment resulted in 28% assignment to therapy only, 25% to cash only, 25% to both, and 22% to neither (see Table A.1). The excess therapy assignments are in part due to chance, and is in part driven by two blocks where excess treatment chips were accidentally used. All regressions include block fixed effects to account for this.

Assignment to the four arms was largely balanced along covariates. Blattman et al. (2017) reported tests of randomization balance for the full sample. That analysis found that, of 57 covariates over three treatments, 14 (8.2%) had a difference with $p < .05$, and within treatment arms the covariates were not jointly significant. Here, in Table A.2, we reproduce these baseline summary statistics and tests of balance, but we limit the sample

to the 833 subjects interviewed at the 10-year endline. We do so to confirm that there is no imbalance introduced by attrition. Column 1 reports the sample mean for each covariate, and Columns 2 to 7 report the coefficients and p values on treatment indicators from ordinary least squares (OLS) regressions of each baseline covariate on three treatment indicators (one for assignment to each treatment arm), controlling for block fixed effects. Column 8 reports the p -value from a joint test of significance of the three coefficients.

Overall, there is minor imbalance. Of 171 coefficients (57 covariates and 3 treatment arms), 12 (7%) have a $p < .05$, and other 7 (4.1%) have a $p < .1$. When we look at tests of joint significance across all arms, 7 of 57 covariates (12%) have a $p < .1$, while 4 (7%) have a $p < .05$. We control for these baseline covariates in all treatment effects regressions in the paper to account for these modest differences.

A.4 Data and attrition

Each survey round was roughly 90 minutes long and delivered verbally by enumerators in Liberian English on handheld tablets. To measure time and risk preferences the respondents also conducted 45 minutes of incentivized games and tests.

We conducted the 10-year survey round roughly 123 months after the completion of Phase 1 treatment, 117 months after Phase 2 treatment, and 109 months after Phase 3. The weighted average gives us a roughly 9.5-year follow-up, which we round to 10 years.

Table A.3 reports the correlates of attrition from a multivariate regression including indicators for each treatment arm and a selection of baseline covariates. Looking at all sources of attrition, in column 2, we see roughly 4 percentage point higher rates of attrition in the Therapy Only (not statistically significant). We see 4 percentage point lower attrition in the Therapy and Cash arm, which is also not significant.

Excess attrition in the Therapy Only arm is primarily due to higher levels of mortality (Column 3, not statistically significant). The dependent variable is an indicator for having died between baseline and the first round of the 10-year survey. Note that point estimates in column 3 differ slightly from Table 1 because we are using a more limited set of baseline covariates for illustrative purposes. We explore patterns and sources of mortality further in Appendix A.1.

Lower attrition in the Therapy+Cash arm was primarily due to better success rate at finding survivors (also significant at the 10% level, in column 5). Otherwise, attrition is not particularly systematic. Looking at Column 2 of Table A.3, for instance, few covariates have large or statistically significant effects on attrition.

Table A.4 reports the results of a single regression across four columns. The regression

Table A.2: Summary statistics and randomization balance, 10-year surveyed sample only

Baseline covariate	Sample Mean (1)	Test of randomization balance(N=833)						F-Test p value (8)
		Assigned Therapy		Assigned Cash		Assigned Both		
		Coef. (2)	p value (3)	Coef. (4)	p value (5)	Coef. (6)	p value (7)	
Age	25.235	-0.136	0.790	0.310	0.556	-0.133	0.795	0.793
Married or partnered, binary	0.149	-0.022	0.572	-0.038	0.335	-0.024	0.527	0.816
# of partners	0.531	0.057	0.422	0.131	0.069	-0.025	0.690	0.111
# of children<15 in household	2.212	-0.622	0.073	-0.768	0.028	-0.643	0.054	0.142
Sees family often	2.345	0.122	0.227	0.286	0.006	0.108	0.295	0.049
Muslim	0.080	0.001	0.965	0.006	0.839	0.018	0.522	0.914
Years of schooling	7.892	0.011	0.974	0.272	0.414	-0.177	0.605	0.572
Currently in school	0.060	-0.018	0.455	-0.019	0.437	-0.010	0.691	0.865
Literacy index (0-2)	1.247	0.116	0.223	0.101	0.282	-0.043	0.661	0.212
Math score (0-5)	2.812	-0.069	0.683	0.055	0.738	-0.370	0.029	0.037
Health index (0-6)	4.852	-0.080	0.571	-0.191	0.200	0.023	0.878	0.482
Has any disabilities	0.068	0.033	0.183	-0.011	0.671	0.000	0.987	0.337
Depression index (0-17)	7.078	0.247	0.437	0.014	0.967	-0.028	0.933	0.788
Distress index (0-21)	7.414	-0.231	0.578	-0.202	0.640	-0.641	0.145	0.502
Relations to commanders index (0-4)	0.427	-0.014	0.862	0.085	0.278	0.047	0.565	0.552
Ex-combatant	0.372	0.068	0.173	0.075	0.143	0.090	0.071	0.294
War experiences index (0-12)	5.785	0.385	0.205	-0.017	0.957	-0.079	0.794	0.328
Weekly cash earnings (USD)	16.615	-4.747	0.041	-5.752	0.013	-3.687	0.116	0.088
Summary index of income, z-score	0.005	-0.196	0.061	-0.097	0.380	-0.107	0.332	0.305
Homeless in past two weeks	0.228	-0.019	0.669	0.020	0.641	-0.003	0.944	0.834
# of days slept hungry, last 7 days	1.248	0.220	0.123	0.328	0.022	0.159	0.244	0.143
Savings stock (USD)	32.661	-11.866	0.101	-13.101	0.067	-10.702	0.162	0.307
Can get loan of 50 USD	0.502	-0.033	0.520	-0.057	0.257	-0.037	0.469	0.725
Can get loan of 300 USD	0.103	-0.016	0.596	-0.002	0.946	0.006	0.857	0.911
Hours/week in illicit activities	13.185	-0.989	0.733	-1.252	0.666	0.168	0.958	0.950
Hours/week in agriculture	0.402	0.414	0.359	-0.152	0.359	0.337	0.451	0.363
Hours/week in low-skill wage labor	18.971	-1.299	0.655	-1.454	0.615	-0.031	0.992	0.924
Hours/week in low-skill business	12.229	0.930	0.706	-0.126	0.960	3.897	0.119	0.342
Hours/week in high-skill work	1.586	-0.081	0.912	0.653	0.376	0.989	0.252	0.594
Years of experience in agriculture	0.742	-0.023	0.913	-0.205	0.336	-0.217	0.318	0.625
Years experience in non-agricultural business	3.050	-0.175	0.718	-0.759	0.119	-0.338	0.477	0.428
Years experience in high-skill work	0.986	-0.422	0.105	-0.576	0.081	-0.037	0.893	0.130
Sells drugs	0.196	-0.009	0.826	-0.018	0.675	0.005	0.907	0.944
Drinks alcohol	0.760	0.073	0.114	0.054	0.234	0.047	0.310	0.456
Uses marijuana	0.582	0.104	0.039	0.077	0.122	0.037	0.469	0.169
Uses marijuana daily	0.431	0.057	0.255	0.029	0.570	0.006	0.912	0.624
Use hard drugs	0.242	-0.026	0.554	0.018	0.692	-0.013	0.784	0.764
Uses hard drugs daily	0.132	-0.074	0.022	0.030	0.429	-0.023	0.509	0.007
Committed theft/robbery in past 2 weeks	0.523	0.034	0.520	0.011	0.826	0.020	0.697	0.932
Number of nonviolent stealing incidents	4.823	-0.584	0.540	-0.812	0.378	-0.443	0.676	0.841
Number of felony stealing incidents	0.429	-0.111	0.532	0.050	0.801	0.009	0.966	0.865
Disputes and fights in past 2 weeks (0-9)	2.077	-0.461	0.415	0.143	0.809	-0.532	0.357	0.429
Aggressive behaviors (mean of 19), z-score	-0.017	-0.032	0.779	0.071	0.510	-0.163	0.133	0.125
Conscientiousness index (0-24)	15.369	-0.051	0.836	-0.255	0.285	-0.270	0.277	0.568
Neuroticism index (0-21)	12.006	-0.060	0.828	0.201	0.453	0.245	0.366	0.513
Grit index (0-21)	13.812	0.172	0.421	0.029	0.893	0.017	0.935	0.820
Reward responsiveness index (0-24)	14.675	-0.204	0.491	0.135	0.657	-0.384	0.203	0.326
Locus of control index (0-24)	14.465	-0.138	0.520	-0.456	0.040	-0.068	0.761	0.171
Impulsiveness index (0-21)	9.324	0.450	0.261	0.178	0.658	-0.288	0.476	0.203
Self esteem index (0-24)	13.541	-0.009	0.974	0.026	0.924	0.126	0.610	0.929
Patience in game play index (0-6)	4.177	0.084	0.675	-0.211	0.303	-0.140	0.508	0.397
Time inconsistency in game play index (0-6)	3.280	-0.211	0.034	-0.024	0.811	-0.119	0.229	0.125
Risk aversion index (0-3)	1.580	-0.050	0.684	-0.018	0.879	0.065	0.598	0.773
Self-reported patience (mean of 7), z-score	0.009	-0.013	0.908	-0.081	0.430	0.030	0.782	0.700
Declared Risk Appetite (mean of 6), z-score	-0.005	0.025	0.818	-0.012	0.910	-0.139	0.213	0.385
Cognitive ability (z-score)	0.046	0.095	0.371	0.135	0.207	-0.080	0.437	0.089
Executive function (z-score)	-0.002	0.067	0.521	0.078	0.457	-0.145	0.184	0.107

Notes: The table reports ordinary least squares regressions of each baseline covariate on treatment indicators, controlling for randomization block fixed effects. We limit the analysis to the 833 members of the sample interviewed at the 10-year endline. All p-values are heteroskedastic-robust, with $p < 0.05$ in bold.

Table A.3: Attrition balance by treatment arm and baseline covariates

	Sample Mean (SD)	Did not survey Estimate (SE)	Deceased Estimate (SE)	Imprisoned Estimate (SE)	Did not survey (Excl. deceased & imprisoned) Estimate (SE)
	(1)	(2)	(3)	(4)	(5)
Therapy Only	0.280 (0.449)	0.044 (0.036)	0.040 (0.029)	0.003 (0.008)	0.001 (0.022)
Cash Only	0.250 (0.433)	-0.018 (0.036)	0.018 (0.029)	-0.009 (0.007)	-0.027 (0.022)
Therapy + Cash	0.249 (0.433)	-0.039 (0.034)	-0.004 (0.027)	0.002 (0.008)	-0.038* (0.021)
Age	25.400 (4.858)	0.005* (0.003)	0.007*** (0.002)	0.000 (0.001)	-0.001 (0.002)
Married or partnered, binary	0.158 (0.365)	0.011 (0.040)	0.011 (0.034)	-0.001 (0.006)	0.001 (0.024)
# of children<15 in household	2.209 (3.174)	-0.002 (0.004)	0.003 (0.004)	-0.001 (0.001)	-0.004* (0.002)
Years of schooling	7.719 (3.287)	-0.009* (0.005)	-0.005 (0.004)	-0.001 (0.001)	-0.003 (0.004)
Cognitive skill index, z-score	0.000 (1.000)	-0.005 (0.016)	-0.010 (0.012)	0.002 (0.004)	0.002 (0.012)
Health index, z-score	-0.004 (1.002)	0.002 (0.012)	-0.004 (0.010)	0.002 (0.002)	0.003 (0.008)
Depression and distress index, std.	0.000 (1.000)	-0.007 (0.013)	-0.008 (0.012)	0.001 (0.002)	-0.000 (0.008)
War experiences index (0-12)	5.859 (2.873)	0.002 (0.005)	0.007* (0.004)	0.000 (0.001)	-0.005 (0.003)
Index of economic success, z-score	-0.003 (0.999)	0.009 (0.014)	0.003 (0.011)	0.000 (0.003)	0.006 (0.009)
Savings stock (USD)	33.753 (67.388)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Hours/week in illicit activities	13.557 (27.253)	0.001 (0.001)	0.001* (0.001)	0.000 (0.000)	-0.001 (0.000)
Hours/week working	45.974 (43.231)	-0.001* (0.000)	-0.001* (0.000)	-0.000 (0.000)	-0.000 (0.000)
Index of antisocial behaviors, z-score	0.001 (1.003)	0.015 (0.016)	0.002 (0.014)	-0.001 (0.005)	0.014 (0.010)
Index of personality measures, z-score	0.000 (0.949)	-0.004 (0.013)	-0.004 (0.011)	-0.002 (0.002)	0.003 (0.008)
Index of patience, z-score	-0.003 (0.993)	-0.033*** (0.013)	-0.030*** (0.011)	-0.000 (0.002)	-0.003 (0.008)
Dep. var. mean		0.166	0.103	0.007	0.056
N		999	999	999	999

Notes: This table regresses different attrition outcomes (not surveyed, deceased, imprisoned, adjusted not surveyed) on the different treatment arms and a subset of the baseline variables. Deceased is an indicator for a reported death before the 10-year survey. Imprisoned is an indicator for being unable to interview the subject at all at the 10-years due to lengthy incarceration.

interacts each covariate with indicators for each treatment arm, to test whether the determinants of attrition vary by arm. Attrition in the control arm is greater among men with more war experiences and lower baseline economic performance, but these are generally reversed in the treatment arms where there is little evidence of selective attrition.

Table A.4: Attrition balance interacting treatment arm and baseline covariates

	Sample Mean	Control	Interaction Coeff.			Differences	
			Therapy Only	Cash Only	Both	Therapy Only vs Both	Cash Only vs Both
		Estimate (SE) [p-value]	Estimate (SE) [p-value]	Estimate (SE) [p-value]	Estimate (SE) [p-value]	Estimate [p-value]	Estimate [p-value]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Levels			0.195 (0.234) [0.404]	0.297 (0.234) [0.204]	0.128 (0.223) [0.566]	-0.067 (0.766)	-0.169 (0.426)
Age	25.400 (4.858)	0.007 (0.007) [0.350]	0.003 (0.009) [0.730]	-0.010 (0.009) [0.231]	-0.001 (0.009) [0.916]	-0.004 (0.766)	0.009 (0.426)
Married or partnered, binary	0.158 (0.365)	0.024 (0.087) [0.780]	-0.078 (0.117) [0.506]	0.061 (0.120) [0.612]	-0.020 (0.114) [0.861]	0.058 (0.766)	-0.081 (0.426)
# of children < 15 in household	2.209 (3.174)	-0.008 (0.007) [0.256]	0.002 (0.010) [0.854]	0.015 (0.011) [0.175]	0.009 (0.011) [0.420]	0.007 (0.766)	-0.006 (0.426)
Years of schooling	7.719 (3.287)	-0.000 (0.010) [0.998]	-0.014 (0.014) [0.326]	-0.015 (0.014) [0.295]	-0.007 (0.013) [0.611]	0.007 (0.766)	0.008 (0.426)
Cognitive skill index, z-score	0.000 (1.000)	-0.028 (0.034) [0.410]	0.035 (0.047) [0.461]	0.008 (0.045) [0.863]	0.043 (0.045) [0.332]	0.009 (0.766)	0.036 (0.426)
Health index, z-score	-0.004 (1.002)	0.019 (0.027) [0.474]	-0.028 (0.038) [0.460]	-0.004 (0.034) [0.906]	-0.021 (0.034) [0.531]	0.007 (0.766)	-0.017 (0.426)
Depression and distress index, std.	0.000 (1.000)	-0.042 (0.028) [0.128]	0.013 (0.042) [0.756]	0.036 (0.036) [0.317]	0.080** (0.041) [0.049]	0.067 (0.766)	0.045 (0.426)
War experiences index (0-12)	5.859 (2.873)	0.016 (0.010) [0.120]	-0.027* (0.015) [0.069]	0.002 (0.014) [0.884]	-0.029** (0.013) [0.029]	-0.002 (0.766)	-0.031 (0.426)
Index of economic success, z-score	-0.003 (0.999)	-0.024 (0.025) [0.342]	0.074* (0.041) [0.074]	-0.014 (0.036) [0.703]	0.076** (0.037) [0.042]	0.002 (0.766)	0.090 (0.426)
Savings stock (USD)	33.753 (67.388)	0.000 (0.000) [0.903]	-0.000 (0.001) [0.984]	0.000 (0.001) [0.436]	0.000 (0.001) [0.643]	0.000 (0.766)	-0.000 (0.426)
Hours/week in illicit activities	13.557 (27.253)	0.001 (0.001) [0.529]	0.001 (0.002) [0.602]	-0.000 (0.002) [0.848]	-0.000 (0.002) [0.772]	-0.001 (0.766)	-0.000 (0.426)
Hours/week working	45.974 (43.231)	-0.001* (0.001) [0.094]	0.001 (0.001) [0.596]	-0.000 (0.001) [0.988]	0.001 (0.001) [0.290]	0.000 (0.766)	0.001 (0.426)
Index of antisocial behaviors, z-score	0.001 (1.003)	-0.007 (0.030) [0.818]	0.013 (0.045) [0.769]	0.039 (0.042) [0.351]	0.029 (0.042) [0.491]	0.016 (0.766)	-0.010 (0.426)
Index of personality measures, z-score	0.000 (0.949)	-0.006 (0.029) [0.826]	-0.015 (0.040) [0.698]	0.009 (0.038) [0.810]	0.008 (0.040) [0.839]	0.024 (0.766)	-0.001 (0.426)
Index of patience, z-score	-0.003 (0.993)	-0.043 (0.028) [0.124]	-0.006 (0.041) [0.887]	0.046 (0.037) [0.216]	0.025 (0.033) [0.460]	0.031 (0.766)	-0.022 (0.426)
Joint F-Test		.413	.784	.801	.264		

Notes: This table regresses the main attrition outcomes (not surveyed) on a subset of the baseline variables, while also interacting them with each treatment arm. This allows the baseline covariate to affect attrition differently on each treatment arm. Column (2) displays the coefficients from the non-interacted control variable, while Columns (3) - (5) display the interacted term for with the Therapy Only, Cash Only and Therapy+Cash treatment arm, respectively. The first row does not display an interaction estimates, but rather the estimates from the corresponding treatment indicator variable. Columns (6) and (7) test the statistical significance of the difference between the Therapy+Cash treatment arm and the Therapy Only and Cash Only, respectively. Finally, the last row presents a joint F-test of the column estimates, excluding the level estimates.

B Additional results

B.1 Empirical specification

We estimate intent-to-treat (ITT) effects via the ordinary least squares regression:

$$Y_{ij} = \tau_1 \textit{TherapyOnly}_i + \tau_2 \textit{CashOnly}_i + \tau_3 \textit{Therapy\&Cash}_i + \mathbf{X}_i \lambda + \gamma_j + \varepsilon_{ij} \quad (2)$$

where *TherapyOnly*, *CashOnly*, and *Therapy&Cash* are indicators for random assignment to the 3 treatment arms. We control for a vector of baseline characteristics, X , and fixed effects for each of the j randomization blocks, γ_j . Y_{ij} is the average outcome from the two proximate survey rounds. To reduce sensitivity to outliers, we top-code continuous variables at the 99th percentile. We estimate heteroskedastic robust standard errors.

B.2 Robustness tests

Appendix Table B.1 reports treatment effects from alternative models and outcomes in 7 columns:

1. The main results on from Table 1
2. Removing the top-coding of extreme values in the outcome components
3. Using an alternative, briefer vector of controls
4. Removing all control variables except for randomization strata fixed effects
5. Using a control vector selected using the double lasso (machine learning) method
6. Instead of one observation per respondent, averaged across the two 10-year survey rounds, we use two observations per respondent and cluster standard errors at the individual level
7. The antisocial behaviors index is constructed using covariance weights rather than equally weighting the component measures

Table B.1: Program impacts on antisocial behaviors index after 10 years: Robustness to different covariates and alternative index construction

	Main	No top-coding of ASB components	Subset of main control vars.	Fixed-Effects Only	Double lasso control vars.	2 obs. per individual	Covariance weighted index
	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Therapy Only	-0.200* (0.106)	-0.204* (0.106)	-0.176* (0.103)	-0.182 (0.115)	-0.205* (0.107)	-0.157** (0.077)	-0.186* (0.106)
Cash Only	-0.075 (0.105)	-0.081 (0.104)	-0.069 (0.103)	-0.067 (0.114)	-0.095 (0.104)	-0.062 (0.078)	-0.087 (0.105)
Therapy + Cash	-0.245** (0.104)	-0.251** (0.104)	-0.256** (0.110)	-0.230** (0.115)	-0.234** (0.104)	-0.204*** (0.075)	-0.247** (0.106)

Notes: The table reports different robustness checks to the intent-to-treat estimates. Column (1) reports the treatment estimates of the main specification. Column (2) runs the main specification with a subset of 15 out of the 57 baseline control variables. Column (3) shows regress a modified version of the outcome variable where the ASB index components that are unbounded have not been top-coded at the 99th percentile. Column (4) shows robustness to eliminating randomization strata fixed-effects. Column (5) shows robustness to including exclusively such fixed-effects. Column (6) allows a double lasso method to choose a subset of the baseline covariates. Column (7) treats each subject response as a separate unit of analysis, while clustering standard errors at such level. Column (8) construct the ASB index according to the covariance of the index component.

Extreme attrition scenarios Even though attrition appears to have been relatively un-systematic, mortality and other attrition is large enough that unobserved selection could influence our conclusions. For example, recall that we are slightly more likely to find men in the Therapy+Cash arm at endline. If the men who died or whom we were unable to find had systematically lower rates of antisocial behavior than the ones we found, then our estimated treatment effects would overestimate the effects of Therapy+Cash.

Table B.2 reproduces our main result in Column 2 and models three selective attrition scenarios in Columns 3 through 5. In Columns 3 and 4 we impute extreme values for missing subjects. That is, we impute systematically “good” outcomes for unfound control group members and “bad” outcomes for unfound treatment group members. For missing control group members, we impute the mean observed outcome plus a 0.1 or 0.2 standard deviation improvement. We do the opposite for missing treatment group members. Thus, we re-estimate treatment effects on the full baseline sample of 999, rather than the observed sample of 833. This is a rather extreme test, as it mechanically creates a 0.2 or 0.4 standard deviation gap in performance between missing treatment and control group members, attenuating our treatment effect estimates by construction.

For Therapy+Cash, we can see that the estimated treatment effect on antisocial behavior survives a gap of 0.4 (and indeed larger). This goes to show the extreme level of systematic

Table B.2: Attrition bound estimates for antisocial behaviors

	Main reg (N=833)	Impute .1 SD (N=999)	Impute .2 SD (N=999)	Trim (N = 393, 384, 383)
	Estimate (SE) [p-value]	Estimate (SE) [p-value]	Estimate (SE) [p-value]	Estimate (SE) [p-value]
	(1)	(2)	(3)	(4)
Therapy Only	-0.200* (0.106) [0.059]	-0.155* (0.090) [0.085]	-0.118 (0.090) [0.189]	0.020 (0.093) [0.830]
Cash Only	-0.075 (0.105) [0.477]	-0.048 (0.088) [0.590]	-0.015 (0.089) [0.870]	-0.048 (0.122) [0.692]
Therapy + Cash	-0.245** (0.104) [0.019]	-0.206** (0.094) [0.028]	-0.181* (0.094) [0.053]	-0.239* (0.123) [0.053]

Notes: The table reports the intent-to-treat estimates on antisocial behaviors for each treatment arm when using different sensitivity analyses to model systematic attrition. Column (2) displays the main results, column (3) shows a regression where the mean + 0.1 standard deviations (SD) is imputed for each missing outcome value in the control group, while a value equal to mean - 0.1 SD is imputed for the other arms. Column (4) repeats this process with a value of +/- 0.2 SD instead. Finally, column (5) implements Lee bounds—trimming (dropping) values as to artificially reduce the imbalance in missing values relative to the control group. The column summarizes the estimates from 3 different regressions.

attrition that would have to be present to change our conclusions. Impacts on the Therapy Only arm, however, are more sensitive to these attrition scenarios, partly because of the lower effect size but mostly because of the higher levels of attrition in that arm.

Finally, we report a trimming exercise (Lee bounds) in Column 5. This approach recalculates treatment effects after dropping (rather than adding and imputing) observations, to equalize imbalance across arms. For instance, the Therapy+Cash arm has roughly 4 percentage points higher response rates than the control arm. The trimming exercise drops the highest-performing members of the Therapy+Cash arm until the two arms have equivalent attrition levels, then recalculates treatment effects.¹¹

Our analysis echoes the extreme values result: the Therapy+Cash impacts are highly robust to dropping the most anti-social members of that arm (to bring attrition to the same levels as the control arm) but the Therapy Only arm is not at all robust to dropping the least antisocial members of the control arm (to bring attrition levels to equal levels).

B.3 Heterogeneity

Table B.3 reports the same intent-to-treat estimates of each treatment arm after 10 years as in Table 1, interacting each treatment indicator with a continuous, standardized measure of baseline antisocial behavior (ASB) with zero mean and unit standard deviation. Heteroskedastic-robust standard errors are reported in parentheses and p-values in brackets.

¹¹Because we have more than one treatment, and because we have different levels of attrition in each arm (sometimes in opposite directions) we calculate the estimates in this column from three separate regressions. In each regression, we restrict the analysis to the control group and the treatment arm in question (i.e., about half the sample), in each case trimming the arm with lower attrition in the direction that would diminish treatment effects.

Table B.3: Heterogeneity in program impacts by baseline antisocial behavior (continuous, standardized)

	Sample Mean (SD)	Antisocial behavior Estimate (SE) [p-value]	Economic variables Estimate (SE) [p-value]	Identity Estimate (SE) [p-value]	Time preferences Estimate (SE) [p-value]	Self-control Skills Estimate (SE) [p-value]	Mental health Estimate (SE) [p-value]	Social networks Estimate (SE) [p-value]	Substance abuse Estimate (SE) [p-value]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Therapy Only	0.266 (0.442)	-0.176* (0.102) [0.083]	0.042 (0.113) [0.712]	-0.040 (0.101) [0.690]	0.153* (0.092) [0.099]	0.134 (0.100) [0.179]	0.109 (0.100) [0.278]	-0.011 (0.110) [0.920]	0.004 (0.087) [0.965]
Cash Only	0.251 (0.434)	-0.062 (0.102) [0.542]	0.004 (0.112) [0.973]	-0.101 (0.097) [0.296]	0.015 (0.091) [0.868]	0.007 (0.100) [0.947]	-0.069 (0.099) [0.484]	-0.040 (0.104) [0.704]	0.016 (0.087) [0.857]
Therapy + Cash	0.264 (0.441)	-0.257** (0.107) [0.017]	0.194* (0.114) [0.090]	0.128 (0.102) [0.211]	0.266*** (0.091) [0.004]	0.104 (0.099) [0.295]	0.211** (0.100) [0.036]	0.110 (0.101) [0.279]	-0.066 (0.092) [0.474]
Therapy Only × Baseline ASB	-0.008 (0.476)	-0.341*** (0.098) [0.001]	0.135 (0.112) [0.226]	0.450*** (0.120) [0.000]	0.296*** (0.094) [0.002]	0.154 (0.094) [0.102]	0.219** (0.098) [0.026]	0.048 (0.121) [0.689]	-0.143* (0.085) [0.092]
Cash Only × Baseline ASB	-0.003 (0.506)	-0.139 (0.145) [0.340]	0.029 (0.104) [0.782]	0.198 (0.127) [0.119]	0.219** (0.091) [0.016]	-0.020 (0.096) [0.836]	0.034 (0.087) [0.702]	0.029 (0.121) [0.813]	-0.151 (0.094) [0.111]
Therapy + Cash × Baseline ASB	-0.010 (0.499)	-0.389*** (0.102) [0.000]	0.207* (0.110) [0.061]	0.305** (0.119) [0.011]	0.205** (0.093) [0.028]	0.147 (0.090) [0.105]	0.211** (0.090) [0.020]	0.145 (0.121) [0.231]	-0.185** (0.089) [0.038]
Baseline ASB	-0.024 (0.975)	0.459*** (0.074) [0.000]	-0.107 (0.085) [0.205]	-0.398*** (0.104) [0.000]	-0.261*** (0.061) [0.000]	-0.140** (0.069) [0.041]	-0.177*** (0.062) [0.005]	-0.082 (0.101) [0.416]	0.250*** (0.061) [0.000]

Notes: The table reports the same intent-to-treat estimates of each treatment arm after 10 years as in Table 1, interacting each treatment indicator with a continuous, standardized measure of baseline antisocial behavior (ASB) with zero mean and unit standard deviation. Heteroskedastic-robust standard errors are reported in parentheses and p-values in brackets.

B.4 Mortality

Mortality was high in our sample. To determine cause of death, we interviewed two friends or relatives of every respondent reported as having died. We collected quantitative and lengthy qualitative explanations on the circumstances around the death, related events, and so on. If there were serious disagreements we sought out additional sources, but in practice this was rare.

The city, which sits in swampy terrain, has endemic malaria among other tropical diseases. Diseases like tuberculosis are also commonplace. The health system is rudimentary, even by regional standards, and the country has one of the lowest human development indexes in the world. Tragically, the city also experienced a major Ebola outbreak in the years between the 1- and 10-year surveys. Changes in antisocial behavior may not be protective under these extreme circumstances.

Table B.4 reports counts of death by reported cause. Tragically, in addition to the 103

deaths that occurred before the first round of the 10-year survey, 6 additional members of the sample died after being interviewed in the first or second round. We report these revised figures in Table B.5.

Our results in Tables 1 and A.3 use an indicator for the 103 deaths, but the results are qualitatively the same if we use an indicator for the 109. In all specifications, the Therapy+Cash arm had similar levels of mortality as the control arm, and the Therapy Only arm had slightly elevated levels of mortality, which are not statistically significant in most specifications. In some specifications, including some using the 109, the impact of Therapy Only on all deaths is statistically significant at the 10 percent level, but there is no meaningful change in the level or significance of the impacts on violent death.

Turning to death by type, almost two-thirds of the deaths are due to illnesses and other health problems—including Ebola, tuberculosis, and a variety of sudden deaths and unexplained illnesses that could be due to any number of (usually) undiagnosed conditions or diseases.

We had no reports of death by drug overdose. Possibly this population is seldom able to obtain or afford a sufficiently large and pure dose to overconsume. But in a small number of illness-related deaths our informants suggested that drug use aggravated an existing illness and contributed to the death.

Finally, we identified 24 violent deaths before the start of the 10 year survey (plus two additional ones in the months following, while the survey was still running in those neighborhoods). Most of these are associated with mob violence—the lynching and killing of a robber by the community. There were roughly equal numbers (4 or 5) in the control, Cash Only and Therapy+Cash arms—equal to about 1.5% of the population of the treatment arm. The highest number of deaths is in the Therapy Only group, with 11 violent deaths before the survey launched, and 13 if we take into account all.

As we saw in Table 1, none of these differences in across arms are statistically significant. Nonetheless, the elevated level of mortality in the Therapy Only group is concerning. Initially, we were concerned that Therapy Only could have motivated young men to leave risky lives and crime, and placed them at risk of crime victimization. But our qualitative investigations suggest not. Most of the violent deaths were mob violence against thieves, and the remainder were often related to disputes within households. Moreover, the higher number of violent deaths in the Therapy Only group is proportionally matched by a rise in deaths due to illness. The elevated levels may simply be statistical noise.

In any event, could excess death and attrition in the Therapy Only arm lead us to overstate the effects of the treatment, by only assessing the behavior of survivors? This is part of a more general concern that selective attrition could influence our results. The

Table B.4: Cause of death for mortality before the 10-year survey

Death type	Control	Therapy	Cash	TP + Cash	Total
Accidental death	0	0	3	1	4
Health complications, drug aggravated	1	2	1	2	6
Health complications, not drug aggravated	13	20	18	11	62
Other	1	4	2	0	7
Violent death	4	11	4	5	24
Total # of deaths	19	37	28	19	103
Death proportion	0.086	0.132	0.112	0.076	0.103

Table B.5: Cause of death for mortality before, during, and after 10-year survey

Death type	Control	Therapy	Cash	TP + Cash	Total
Accidental death	0	0	4	1	5
Health complications, drug aggravated	1	3	1	2	7
Health complications, not drug aggravated	13	21	18	12	64
Other	1	4	2	0	7
Violent death	4	13	4	5	26
Total # of deaths	19	41	29	20	109
Death proportion	0.086	0.146	0.116	0.080	0.109

extreme attrition analysis above suggests that this is not the case. Rather, the results are highly robust to imputing extreme antisocial behavior for missing sample members.

B.5 Secondary outcomes and potential mechanisms

Measurement

Finally, here we elaborate on the components of the 7 secondary outcomes. Tables B.6 to B.12.

Table B.6: Program impacts on components of the time preferences index

	N	Control Mean	Average Treatment Effects								
			Therapy Only			Cash Only			Both		
			β	SE	p-values	β	SE	p-values	β	SE	p-values
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	
Forward-looking time preferences	832	-0.384	0.132	0.096	0.168	0.013	0.097	0.890	0.247***	0.093	0.008
Self-reported svy. questions on patience (3 Q's, z-score)	832	-0.226	0.082	0.095	0.392	-0.100	0.095	0.291	0.177*	0.095	0.062
Self-reported svy. questions on time inconsistency (3 Q's z-score)	832	-0.527	-0.079	0.112	0.484	-0.131	0.113	0.247	0.050	0.104	0.631
Variables obtained from patience games											
Incentivized trade-offs (0 to 6)	828	4.149	0.019	0.179	0.916	0.047	0.184	0.798	0.285	0.184	0.122
Hypothetical trade-offs (0 to 6)	832	3.703	0.497**	0.241	0.040	0.321	0.242	0.184	0.627***	0.238	0.009
Hypothetical discount rate (.9 to 4)	828	2.005	-0.040	0.120	0.741	-0.055	0.117	0.643	-0.212*	0.117	0.071
Variables obtained from time inconsistency games											
Incentivized trade-offs (-3 to 3)	828	0.243	-0.057	0.077	0.459	-0.073	0.075	0.329	-0.029	0.074	0.699
Hypothetical trade-offs (-3 to 3)	832	0.121	-0.180*	0.092	0.050	-0.053	0.096	0.580	-0.093	0.096	0.330
Hypothetical discount rate (-.31 to 3.1)	828	0.005	0.025	0.092	0.789	0.079	0.096	0.413	0.020	0.088	0.825

Future orientation Becoming more self controlled and forward-looking are central components of many behavior change programs (Almlund et al., 2011; Borghans et al., 2008). We attempted to measure discount rates and time inconsistency (corresponding to β and δ preferences) in four ways:

- *Incentivized trade-offs.* Following the survey, subjects were asked to play a set of “real money games” where they had to make a series of intertemporal choices between money at one point in time versus more money later in time, with some probability of a payout. The average payout was about \$3, roughly a day’s wages. Based on game play, we assigned present and future patience scores for each respondent.
- *Hypothetical trade-offs.* During the survey questionnaire, well before the incentivized games, we asked respondents to make the exact same series of trade-offs as above, but in a purely hypothetical setting. We constructed the patience and time inconsistency proxies in exactly the same manner.
- *Hypothetical discount rate.* We also attempted to measure the discount rate in a second way. As in Holt and Laury (2002), we asked respondents a series of hypothetical intertemporal choices for larger amounts of money (on the order of US\$10-30, about a week’s wages) and calculated discount rates based on each respondent’s first switch from a present preference to a future preference.
- *Self-reported survey questions.* We asked respondents six qualitative questions to gauge their self-reported levels of patience and time inconsistency. For example, respondents were asked to place themselves on a ladder from 0 (least patient) to 5 (most patient) as one measure of self-reported patience, and how much they agree with statements such as “When I get money, I spend it quickly”.

Table B.7: Program impacts on components of the self control index

	N	Control Mean	Average Treatment Effects								
			Therapy Only			Cash Only			Both		
			β	SE	p-values	β	SE	p-values	β	SE	p-values
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	
Self-control skills	832	-0.110	0.178*	0.103	0.086	0.049	0.103	0.633	0.119	0.100	0.235
Impulsiveness (z-score)	829	0.072	-0.054	0.105	0.605	0.001	0.105	0.995	-0.068	0.099	0.490
Conscientiousness (z-score)	832	-0.054	0.110	0.114	0.336	-0.035	0.108	0.748	0.026	0.111	0.814
Perseverance / GRIT (z-score)	832	-0.058	0.145	0.106	0.172	-0.057	0.105	0.587	0.042	0.106	0.688
Reward responsiveness (z-score)	832	0.083	-0.118	0.108	0.275	-0.215**	0.109	0.049	-0.150	0.106	0.158

Self control We also measured skills of self control by adapting to the local context several standard psychometric questionnaires for four constructs that psychologists associate with less impulsive and more planful behavior:

- *Impulsiveness*, the inability to control thoughts and actions, using 9 questions from the Barrett Impulsiveness Scale;
- *Conscientiousness*, the tendency to be self-disciplined and purposeful, using 8 questions from the NEO-five factor personality inventory (Costa Jr. and McCrae, 1997);
- *Grit*, the ability to press on in the face of difficulty, using 7 questions on perseverance from the GRIT scale (Duckworth and Quinn, 2009); and
- *Reward responsiveness*, whether people are motivated by immediate (typically emotional) rewards, from the Behavioral Inhibition/ Behavioral Activation Scale (Robinson and Berridge, 2000).

Table B.8: Program impacts on components of the identity and values index

	N	Control Mean	Average Treatment Effects								
			Therapy Only			Cash Only			Both		
			β	SE	p-values	β	SE	p-values	β	SE	p-values
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	
Identity and values	832	-0.017	-0.038	0.101	0.703	-0.092	0.099	0.349	0.100	0.099	0.310
Attitudes toward use of violence (sum of 11 indicator Q's.)	832	1.060	0.017	0.152	0.910	-0.032	0.149	0.830	0.012	0.151	0.936
Attitudes toward criminality, (sum of 12 indicator Q's.)	832	2.984	-0.181	0.181	0.318	-0.119	0.165	0.471	-0.232	0.172	0.177
Attitudes toward political violence (sum of 4 indicator Q's.)	794	0.178	0.029	0.086	0.741	0.183*	0.095	0.054	-0.015	0.081	0.853
Index of appearance (z-score)	828	0.161	-0.079	0.085	0.354	-0.112	0.091	0.218	0.026	0.084	0.756
Prosocial behavior (z-score)	832	0.187	-0.068	0.115	0.552	-0.216*	0.114	0.060	-0.034	0.113	0.761

Anticriminal identity and values Research in both psychology and economics supports the idea that groups have well-defined norms of behavior, and that people receive emotional benefits from acting in accordance with the norms of their perceived group (Almlund et al., 2011; Shayo, 2020). To some extent people may also be able to change their perceived social category, and with it values that reward and penalize certain behaviors (Akerlof and Kranton, 2000). Related, criminologists sometimes refer to a similar process of “knifing off” from old social rules and behaviors, and associate these changes with significant turning points in life, such as marriage, a move, or a life-threatening experience (Maruna and Roy, 2007). This literature ties successful knifing off to having a new “script” for the future. The STYL program is intended to be that script.¹²

¹²There are also parallels between STYL and socialization into military groups, street culture, gangs and armed groups. Such groups use similar techniques (appearance change, practice, modeling) to shape young men’s social identity and behavior (Vigil, 2003; Wood, 2008; Maruna and Roy, 2007). NEPI designed STYL to reverse this process.

To assess this, we first attempted to measure values directly, using a set of 33 self-reported attitudes towards the appropriate use of crime and violence in the men’s own lives—indicators of the degree to which they had internalized mainstream social norms. Second, we measured a range of prosocial behaviors, including group memberships, group and community leadership, and contributions to local public goods. Finally, the therapy encouraged men to change their appearance as part of the identity change, and we asked survey enumerators to record their subjective impressions: quality of dress, shoes, cleanliness, and smell.

Table B.9: Program impacts on components of the mental health index

	N	Control Mean	Average Treatment Effects								
			Therapy Only			Cash Only			Both		
			β	SE	p-values	β	SE	p-values	β	SE	p-values
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	
Positive self-regard/mental health	828	-0.089	0.088	0.104	0.399	-0.032	0.102	0.754	0.207**	0.102	0.042
Neuroticism (<i>z</i> -score)	806	0.054	-0.101	0.112	0.364	0.013	0.107	0.902	-0.107	0.112	0.340
Locus of control (<i>z</i> -score)	806	0.069	-0.033	0.114	0.776	-0.249**	0.114	0.029	-0.094	0.112	0.399
Self esteem (<i>z</i> -score)	806	-0.027	-0.067	0.110	0.539	-0.042	0.106	0.693	0.144	0.107	0.178
Summary index of subjective well being (<i>z</i> -score)	828	-0.055	0.101	0.083	0.223	-0.010	0.084	0.907	0.347***	0.086	0.000
Depression, 6 Q’s. (0 to 18)	802	7.989	-0.384	0.357	0.283	-0.251	0.361	0.487	-0.595*	0.360	0.099
Distress (<i>z</i> -score)	802	0.355	-0.155	0.110	0.157	0.011	0.108	0.920	-0.093	0.104	0.369

Mental health This family groups a set of 6 admittedly wide-ranging mental health-related outcomes in the interests of minimizing the number of families.

Half our mental health family index is related to positive self-regard. Poor self-regard has been linked with many aspects of negative behavior and counterproductive or extreme risk-seeking behavior (Coopersmith, 1967). Some research suggests self-regard is captured by an interrelated set of psychological scales, including: (i) neuroticism, the tendency to experience emotional instability or anxiety, assessed with eight questions from the NEO-5 factor personality inventory; (ii) self-esteem; and (iii) locus of control, the extent to which individuals believe they versus fate control events affecting them, measured using eight questions from a standard questionnaire (Judge and Bono, 2001; Sapp and Harrod, 1993). Arguably related to positive self-regard, we also collected a classic happiness measure, asking men to rank their subjective well-being in absolute terms and relative to others in their community.

A second element of the mental health index is depression and distress. We assessed 6 symptoms of depression and 12 symptoms of posttraumatic stress (distress), based on a locally adapted instrument used previously with ex-combatant populations in Liberia (Blattman and Annan, 2016). We group this with positive self-regard as a mental health family in the interests of minimizing the number of families.

Substance abuse The survey collected self-reported use of alcohol, marijuana, or hard drugs, including frequency. We code an indicator for a frequency of usually or daily.

Table B.10: Program impacts on components of the substance abuse index

	N	Control Mean	Average Treatment Effects								
			Therapy Only			Cash Only			Both		
			β	SE	p-values	β	SE	p-values	β	SE	p-values
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	
Substance abuse	832	0.141	-0.052	0.083	0.534	-0.040	0.082	0.623	-0.096	0.083	0.245
Usually drinks (0-1)	832	0.690	-0.065	0.041	0.114	-0.043	0.042	0.312	-0.023	0.041	0.571
Usually uses marijuana (0-1)	832	0.462	-0.038	0.042	0.368	-0.034	0.042	0.421	-0.075*	0.041	0.067
Usually takes hard drugs (0-1)	832	0.220	0.054	0.038	0.159	0.038	0.038	0.314	0.012	0.038	0.759

Table B.11: Program impacts on components of the social networks index

	N	Control Mean	Average Treatment Effects								
			Therapy Only			Cash Only			Both		
			β	SE	p-values	β	SE	p-values	β	SE	p-values
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	
Quality of social networks	802	0.011	-0.026	0.112	0.814	-0.072	0.109	0.513	0.085	0.102	0.407
Peers quality (z-score)	802	0.371	-0.191**	0.092	0.038	-0.139	0.087	0.111	0.027	0.082	0.746
Quality of family relations (z-score)	802	0.042	0.059	0.106	0.577	-0.038	0.104	0.718	0.101	0.103	0.327
Ex-commanders ties (z-score)	791	-0.219	0.229**	0.106	0.032	0.085	0.097	0.382	0.098	0.092	0.286
Big men ties (z-score)	802	0.042	-0.102	0.109	0.351	-0.024	0.109	0.825	-0.048	0.112	0.671

Social networks We measured the traits, positive and negative, of men’s five closest peers.¹³ We also asked about closeness to and support received from family members, former rebel commanders, and “big men” (intended to connote a criminal boss).

Economic performance We measured income in three ways: (i) estimated weekly profits from all activities in each of the two weeks before each survey round; (ii) an index of wealth using durable assets (such as home quality, furnishings, vehicles, and livestock); and (iii) consumption in the two weeks prior to each survey round. We also measured their accumulated stock of savings; their current valuation of their business assets; the number of hours worked across all economic activities in the two weeks prior to each survey round; and an indicator for currently being homeless and sleeping in the streets.

¹³We ask men who their five closest peers are, by name, and then ask whether they hold any of 20 qualities ranging from positive (they work hard, save, go to school) or negative (the steal, do drugs, get in fights).

Table B.12: Program impacts on components of the economic performance index

	Average Treatment Effects										
	N	Control Mean	Therapy Only			Cash Only			Both		
			β	SE	p-values	β	SE	p-values	β	SE	p-values
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	
Economic performance	832	0.069	0.081	0.115	0.483	0.026	0.115	0.823	0.190*	0.112	0.090
Profit 7d avg (USD, capped 99th)	832	29.390	1.649	3.492	0.637	1.876	3.430	0.585	3.328	3.357	0.322
Index of wealth: housing quality and assetss (z-score)	832	0.054	-0.005	0.107	0.962	-0.123	0.108	0.254	0.190	0.116	0.102
Total consumption last 2w (USD)	832	109.600	1.479	10.994	0.893	9.159	12.114	0.450	1.554	12.602	0.902
Savings stock (USD, capped 99th)	832	82.319	-1.644	15.768	0.917	-10.729	15.927	0.501	10.120	16.566	0.541
Business stock (USD, capped 99th)	832	139.825	-17.002	40.023	0.671	22.874	37.675	0.544	21.876	39.039	0.575
# Hours worked past 7d	832	27.903	6.545**	2.657	0.014	3.353	2.493	0.179	4.119	2.572	0.110
Is sleeping in the strees now (indicator)	832	0.132	-0.010	0.029	0.726	0.017	0.029	0.552	-0.032	0.029	0.266

Correlation with primary outcomes

Any mechanism, to be influential, must not only be correlated with treatment but also with the primary outcome. Table B.13 reports these simple bivariate correlations.

Table B.13: Correlation between secondary outcomes and antisocial-behaviors index

	Sample Mean (SD)	Estimate (SE)	Estimate (SE)	Estimate (SE)	Multivariate regression Estimate (SE)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Economic performance	0.112 (1.029)	-0.125*** (0.035)							0.101*** (0.029)
Forward-looking time preferences	-0.247 (0.924)		-0.341*** (0.037)						-0.027 (0.036)
Self-control skills	-0.004 (0.997)			-0.474*** (0.040)					-0.205*** (0.034)
Identity and values	-0.011 (0.994)				-0.521*** (0.043)				-0.341*** (0.047)
Positive self-regard/mental health	0.014 (0.976)					-0.391*** (0.033)			-0.127*** (0.031)
Substance abuse	0.001 (0.995)						0.472*** (0.037)		0.216*** (0.035)
Quality of social networks	-0.000 (1.000)							-0.142*** (0.044)	0.038 (0.037)
Dep. var. mean		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
N		833	833	833	833	829	833	803	803

B.6 Qualitative data validation

Approach to validation We selected and trained eight of the study’s most talented qualitative research staff as validators, all Liberians. In the ten days following the survey, a validator visited the respondent over four days, spending several hours each day in conversation and observation. Validators shadowed respondents as they went about their day, rather than conduct formal interviews. They raised target topics through indirect questions while chatting.

Validators developed techniques to foster trusting relationships and to build rapport: becoming close to street leaders; eating meals with subjects; sharing personal information (including similar acts they or their friends engaged in); and mirroring participants' appearance and vernacular as appropriate. Validators would also observe the respondent's behavior from afar, as well as converse with peers and family. The goal was to attain insider status, and thus reduce the chance of misreporting. The premise was that time, a focus on a small number of behaviors, and trust/rapport building would mean that respondents were less willing, or feel less able, to deceive a more familiar person, who also knows them better. Validator also had the opportunity to clear up misunderstandings and get a more accurate assessment of the behaviors. By discussing sensitive behaviors openly, relating their own experiences and that of friends, validators sought to dispel any notion that certain answers are more desirable, or would result in any strategic gains.

Without knowing the respondent's survey response, y^s , the validators coded an indicator of whether or not the respondent engaged in the behaviors in the two weeks prior to the survey, y^v . The authors reviewed the evidence and the coding for every case. In general, we used a relatively high standard of evidence, only coding $y^v = 1$ for a direct admission of the behavior or persuasive statements that they did not engage in the behavior.

The validators only witnessed or received third-party evidence of the behavior in a fifth of cases, but neither was considered sufficient evidence for a final coding. Both had to be followed by questions confirming that the respondent also engaged in the behavior in the two weeks prior to the survey. In general, we used a relatively high standard of evidence, only coding $y^v = 1$ if the validator directly observed the behavior or the respondent directly admitted it.

If this technique simply reproduced the errors in the survey data, then the validation is little help. The key assumption is that four days of building trust and gathering extensive information, regarding just six behaviors, reduced experimenter demand and other biases correlated with treatment compared to responses during a 300-question, 90-minute questionnaire.

Nonetheless, y^v is not free from error. For instance, the requirement of a direct admission, the disruption in people's lives, errors in recall periods, or increased social desirability bias from scrutiny all undoubtedly led to systematic errors in y^v . These errors, however, are not necessarily correlated with treatment. This is possible, for example, because validators could have learned men's treatment status in conversation, and this could have biased their coding. Nonetheless, we designed the trust-building and evidentiary standards to minimize this risk.

With a sample of 240, we estimate we are powered to detect average under- or over-

reporting of at least 14%, and error correlated with treatment of 28%. Because of power concerns, we pay close attention to the sign, magnitude, and confidence interval for β_1 .