Appendix for online publication

A   Baseline sample

A.1   Full summary statistics and balance tests

Table A.1 expands the balance table in the main paper for the full set of baseline covariates available and used in the treatment effects regressions.

Table A.1: Baseline statistics and balance test

Baseline covariate	Full Sample N=999 Mean	Test of randomization balance			
		Assigned to Therapy		Assigned to Cash	
		Coeff.	p-value	Coeff.	p-value
	(1)	(2)	(3)	(4)	(5)
Age	25.40	-0.13	0.69	0.06	0.82
Married or living with a partner	0.16	0.00	0.98	-0.02	0.36
# of women supported	0.53	-0.04	0.12	0.01	0.61
# children under 15 in household	2.21	-0.14	0.49	-0.11	0.45
How often sees members of family (1-4)	2.37	-0.06	0.28	0.09	0.14
Muslim	0.10	0.03	0.21	0.00	0.90
Years of schooling	7.72	-0.13	0.56	0.03	0.92
Currently in school	0.06	0.00	0.68	-0.01	0.63
Literacy index (0-2)	1.23	0.00	0.91	-0.02	0.58
Math score (0-5)	2.79	-0.17	0.10	-0.06	0.64
Ability to perform daily activities (0-6)	4.87	0.12	0.19	-0.04	0.66
Has any disabilities	0.08	0.01	0.39	-0.02	0.25
Symptoms of depression (0-18)	7.09	0.08	0.52	-0.10	0.46
Symptoms of distress (0-21)	7.46	0.06	0.59	-0.20	0.18
Relationships with ex-commanders, (0-4)	0.45	-0.03	0.56	0.02	0.60
Ex-combatant	0.38	0.01	0.64	0.03	0.34
War experience (0-14)	5.86	-0.01	0.94	-0.18	0.19
Weekly cash earnings (USD)	68.30	2.66	0.29	-7.49	0.22
Index of 38 durable assets, z-score	0.00	-0.06	0.30	0.07	0.50
Currently sleeping on the street	0.24	-0.01	0.42	-0.01	0.38
Times went hungry in past week	1.26	0.04	0.60	0.09	0.26
Savings stock (USD)	33.75	-0.53	0.89	-3.31	0.46
Able to get a loan of $50	0.52	0.00	0.78	-0.03	0.30
Able to get a loan of $300	0.11	0.01	0.75	0.00	0.97
Hours/week in potentially illicit activities	13.55	0.65	0.56	-0.26	0.81
Hours/week in agriculture	0.36	0.41	0.02	-0.06	0.79
Hours/week in low-skill wage labor	19.39	-0.83	0.81	0.10	0.97
Hours/week in in low-skill business	11.53	3.59	0.02	1.83	0.14

Baseline covariate	Full Sample N=999 Mean	Test of randomization balance			
		Assigned to Therapy		Assigned to Cash	
		Coeff.	p-value	Coeff.	p-value
	(1)	(2)	(3)	(4)	(5)
Hours/week in high-skill work	1.51	0.18	0.67	0.85	0.00
Experience in agriculture (years)	0.78	-0.11	0.33	-0.16	0.01
Experience in low skill work (years)	2.96	0.12	0.53	-0.24	0.27
Experience in high skill work (years)	0.96	0.02	0.75	0.06	0.71
Sells drugs	0.20	0.01	0.49	0.00	0.97
Drinks alcohol	0.75	0.05	0.31	0.03	0.44
Uses marijuana ever	0.59	0.03	0.47	0.01	0.59
Uses marijuana daily	0.44	0.03	0.32	0.01	0.57
Uses hard drugs ever	0.26	-0.01	0.62	0.02	0.54
Uses hard drugs daily	0.15	-0.05	0.03	0.03	0.41
Committed theft in past 2 weeks	0.53	0.02	0.68	0.01	0.72
# nonviolent thefts in past 2 weeks	5.10	-0.20	0.57	-0.11	0.90
# violent thefts in past 2 weeks	0.70	-0.14	0.41	-0.05	0.84
# fights and disputes in past 2 weeks	2.16	-0.19	0.46	0.10	0.83
Aggressive and hostile behaviors index, z-score	0.00	-0.05	0.45	0.00	1.00
Conscientiousness index (0-24)	15.33	-0.03	0.62	-0.14	0.26
Neuroticism index (0-24)	12.08	-0.02	0.84	0.24	0.16
Perseverance/GRIT index (0-21)	13.79	0.06	0.59	-0.05	0.79
Reward responsiveness index (0-24)	14.67	-0.23	0.19	-0.06	0.73
Locus of control index (0-24)	14.45	0.13	0.38	-0.17	0.17
Index of 8 self-reported impulsiveness questions (0-24)	9.35	-0.03	0.90	-0.30	0.07
Self esteem index (0-24)	13.47	-0.08	0.56	0.01	0.94
Behavioral game play:					
Patience index (0-6)	4.12	0.11	0.45	-0.05	0.77
Time inconsistency index (0-6)	3.27	-0.18	0.01	0.02	0.74
Risk aversion index (0-3)	1.57	0.06	0.49	0.00	0.97
Declared time preference (z-score)	0.00	0.01	0.82	-0.01	0.84
Declared risk appetite (z-score)	0.00	-0.03	0.82	-0.08	0.07
Executive function (z-score)	0.00	0.02	0.64	0.00	0.96
Cognitive ability (z-score)	0.00	-0.07	0.18	-0.03	0.74
R-squared		0.12		0.05	
p-value on F-statistic on all covariates		0.64		0.90	

A.2 Neighborhoods and recruitment

Table A.2 describes each of the study neighborhoods where we recruited, along with population estimates. We report the estimates of the number of all adult males, as well as our low-end estimates of the number of target males in each neighborhoods—men 18

to 35 in the bottom decile of income.

A.3   Correlates of compliance and attrition

Table lists survey response rates by treatment group and survey wave (short term, pooling 2- and 5-week surveys, and long term, pooling 11- and 13-month surveys). It also reports the p-value from a t-test of the difference between the response rate in each treatment group and the control group. None of the differences are statistically significant, and all are within about a percentage point of the control group response rate. The control group response rate is a tiny bit lower in the long run surveys and a tiny bit higher in the short run ones. But none of these differences control for covariates or even strata fixed effects, as in the next table.

We analyze the correlates of attrition in Columns 1 and 2 of Table A.4. We use a selection of baseline covariates to reduce collinearity and thus aid interpretation. Results with full covariates draw similar conclusions. We define compliance as "attended at least 5 days of therapy", but attendance data was not collected the first week, and so in principle this is an indicator for attending at least 8 days of therapy (about 2.5 weeks).

B   Additional intervention details

After completing the pilot, we decided on a target sample of 1,000. This target was based on maximum program capacity and financial constraints. Based on the pilot, we estimated that the Minimum Detectible Effect for the full 1,000 (for each treatment) would be a 0.12 standard deviation change in a standardized dependent variable for a two-tail hypothesis test with statistical significance of 0.05, statistical power of 0.80, an intra-cluster correlation of 0.25, and the proportion of individual variance explained by covariates as 0.10.

B.1   Randomization protocols

For the therapy and cash randomization, men in each block took turns drawing colored chips from an opaque fabric bag. In general, the bag was shaken and then the subject was instructed to turn away and to place one arm into the bag and to draw out a single chip. The color was confirmed and recorded.

In the cash instance, men were randomized in roughly equal sized blocks of about 50 people. Each man was invited into a private room to draw to ensure privacy and safety.

Table A.2: Recruitment neighborhoods

Phase	Neighborhood	Short description	Estimated males		Recruited		
			All	Target	No.	% all	% target
1	Red Light 1	Peri-urban, along the main road from Monrovia to the northeast of the country, residential but the site of one of the major markets in the city, mixed income neighborhood.	23 422	1 171	100	0.4%	8.5%
2	Red Light 2	Peri-urban, along the main road from Monrovia to the NE of the country, residential but the site of one of the major markets in the city, mixed income neighborhood.	36 434	1 822	219	0.6%	12.0%
	Central Monrovia	The area consists of Mamba Point and West Point, one of the busiest business areas in Monrovia.	32 345	1 617	179	0.6%	11.1%
3	New Kru Town	Peri-urban, the north of Bushrod Island, the transit point to counties of the northwest of Liberia, notorious for petty crime.	28 704	1 435	240	0.8%	16.7%
	Logan Town	Peri-urban, the middle of Bushrod Island, next to Freeport of Monrovia, many garages and small shops and booths.	22 100	1 105	86	0.4%	7.8%
	Clara Town	Peri-urban, the south of Bushrod Island, next to Central Monrovia, lots of car-loaders and wheelbarrowers.	23 921	1 196	175	0.7%	14.6%
All			166 926	8 346	999	0.6%	12.0%

Notes: Total male population estimates come from the authors' calculations based on data from Liberia Institute of Statistics and Geo-Information Services (LISGIS). To get an estimated number of target males we assume half are in the age range of 18-35 and take the bottom 10% decile as our targets. Red Light 1 includes Gorbachop, Woodcamp, Reservoir, Pipeline, Soul Clinic, and Sugar Hill. Red Light 2 includes Turtle Base, Chicken Poultry, Ma Kebbeh Gas Station area, Sugar Hill, Bassa Town, Goba Chop Community, Morris' Farm, Bernard Farm, Pipeline Road, Zayzay Community, Coca Cola Factory Community, Plank Field Community, Banana Bush Community, Soul Clinic Community, and Wood Camp. Central Monrovia includes Mamba Point and West Point areas. New Kru Town also includes part of Calwell. Logan town also includes part of Mamba Point and West Point that are not covered in Phase 2.

Table A.3: Survey response rates by wave and treatment status

	Treatment group				
	Control	Treatment Only	Cash Only	Treatment + Cash	All
Short-Term Response Rate	92.1%	91.0%	89.7%	91.5%	91.0%
p-value vs. control		0.65	0.36	0.83	
Long-Term Response Rate	91.8%	92.9%	94.8%	94.8%	93.6%
p-value vs. control		0.65	0.18	0.18	
All	91.9%	91.9%	92.3%	93.2%	92.4%
p-value vs. control		1.00	0.84	0.48	

This procedure was explained to the entire group, and all chips were placed into the bag in front of everyone. Then the bag was taken into a private room, and participants were called into the room individually. If they wished, they could inspect the bag to confirm that there were still chips of both colors inside. After everyone present had drawn, staff drew the remaining chips for the no-shows.

In the case of therapy, men were randomized each day, according to how many were recruited and surveyed in that neighborhood. This led to blocks ranging in size from 1 to 20, though the vast majority of blocks contained roughly 7 to 15 people. The draw was not as private as the cash draw, and men observed the outcomes of others drawing at the same time. Those who lost in the therapy randomization were offered a free meal along with the opportunity to discuss their situation with someone, and they were transported to a location of their choosing. A small percentage of the men were visibly upset and refused to engage at this point.

B.2 Therapy

The standard STYL curriculum tended to be longer and broader than the two noncognitive skill and value changes that we study. For the purposes of this study, we worked with NEPI to streamline and focus the traditional STYL curriculum in two ways. First, we further grounded the approach in terms of CBT, emphasizing more practice over lectures. In general these modifications were quite modest, since the program already incorporated these techniques. Second, we asked NEPI to exclude modules not relevant to their theories of change: interpersonal skills; conflict resolution skills; dealing with war trauma and PTSD; career counseling; and community leadership.

To clarify and validate NEPI's curriculum, a Liberian qualitative researcher acted as a participant observer throughout one of the two Phase 1 pilot classes. Based on NEPI's

Table A.4: Baseline correlates of survey attrition and treatment compliance for select covariates

	Dependent variable			
	Unfound		Attended >5d of therapy	
Baseline covariate	Coeff.	Std. Err.	Coeff.	Std. Err.
	(1)	(2)	(3)	(4)
Assigned to therapy only	0.004	[0.016]		
Assigned to cash only	-0.004	[0.017]		
Assigned to therapy & cash	0.000	[0.016]		
Age	0.001	[0.001]	0.000	[0.003]
Married or living with partner	-0.017	[0.018]	-0.056	[0.038]
# children under 15 in household	-0.002	[0.002]	0.006	[0.004]*
Years of schooling	0.001	[0.002]	0.019	[0.005]***
Cognitive skills, z-score	0.007	[0.007]	-0.027	[0.016]*
Ability to perform activities of daily life (0-6)	0.001	[0.004]	-0.007	[0.008]
Mental health, z-score	0.017	[0.007]***	-0.014	[0.011]
War experience (0-14)	-0.003	[0.002]	-0.001	[0.004]
Earnings, assets and consumption, z-score	0.011	[0.007]*	-0.016	[0.015]
Savings stock (USD)	0.000	[0.000]	0.000	[0.000]
Hours/week working in potentially illicit activities	0.000	[0.000]	-0.001	[0.001]
Hours/week working (total)	0.000	[0.000]	0.000	[0.000]
Index of all antisocial behaviors, z-score	-0.005	[0.008]	0.034	[0.016]**
Index of personality skills, z-score	0.009	[0.007]	0.029	[0.014]**
Patience index (0-6)	0.000	[0.003]	0.010	[0.007]
Mean of dependent variable	0.073		0.903	
R-squared	0.096		0.298	
P-value for test of joint significance	0.092		0.019	

Notes: Columns (1)–(2) pool all endline survey rounds and report the coefficients and standard errors from an OLS regression of an indicator for attrition on the baseline covariates. Columns (3)–(6) report the coefficients and standard errors from an OLS regression of an indicator for compliance on the same covariates, restricting the samples to people assigned to the respective treatment groups only. Block fixed effects are included in all regressions but are omitted from this table. Robust standard errors are clustered at the block level. *** p<0.01, ** p<0.05, * p<0.1

training materials, our analysis of the theoretical grounding of the therapy, and this participant observation, we and NEPI developed a full program manual for the intervention, available at http://chrisblattman.com/documents/policy/2015.STYL.Program.Manual.pdf. The manual details the history and theory of the interventions, guidelines for recruitment of trainers and participants, training suggestions, the full curriculum, and guidelines for out-of–classroom engagement.

The curriculum has eleven main modules, which we present here with some examples of goals and activities:

1. *Transformation.* A tenet of CBT is that the therapist explicitly sets goals with participants and lays out the therapeutic strategy. This module introduces the concept of transformation, its significance, and the processes involved in transforming oneself.

   - The men are introduced to the techniques that will be used (role playing, lectures, storytelling, etc.), homework assignments, home visits, and the reasons for each.

   - The module also introduces ground rules for behavior, in terms of being respectful, practicing listening, waiting your turn, etc. The men do not necessarily have these skills, or haven't exercised them in some time, and learning to abide by these behavioral rules is an important part of the therapy.

   - Facilitators also begin to teach the songs, slogans, and call-and-response that will be used repeatedly throughout the course. These songs and slogans serve as important reminders of rules of behavior for the men to follow. They also can be used to bring order to a disorderly or inattentive group.

   - There are symbolic rituals to indicate a break in their lives. For example, the men write their "street names" and aliases on sheets of paper and they are burned together.

2. *Substance Abuse.* This module defines substance abuse and discusses its ill effects, as well as steps for moving past it. It explicitly encourages participants to reduce their consumption of drugs, alcohol, and tobacco. They are cautioned against cutting drugs entirely, to avoid withdrawal problems.

   - Men talk through and list reasons that they use drugs. The idea is to make them consciously aware of the reasons for their own behavior and risk factors

in their lives. They also talk through the ill effects. Men talk through publicly about ways in which drugs have adversely impacted their own lives, sharing experiences.

- Men role play situations where they could be pressured to use drugs and practice strategies for saying no.

- An outside speaker comes to the classroom, often a former graduate of the therapy, to talk about their experiences with drugs and what it did to their lives, as well as what strategies they used to emerge. Men discuss strategies they can use in their own lives. They practice some of these as homework and come back to discuss their experiences with the class.

3. *Body Cleanliness.* The module explores the health, psychological, and social benefits of maintaining body cleanliness. Participants are encouraged to change behaviors that alienate them, and to present a public image (such as hair and dress) that promotes positive social interactions with community members.

- Body uncleanliness is defined and highlighted as a problem mainly by getting men to discuss and volunteer their own opinions and experiences in a group.

- The facilitators bring in a hair cutter, an electric shaver, and a set of nail clippers for men to clean up if they like.

4. *Garbage/Dirt Control.* An extension of the previous module, this module highlights the importance of cleanliness in participants' environments, and the ill effects of living in a dirty environment. It aims to help them maintain clean, healthy, and orderly living spaces.

- Facilitators present the men with pictures of dirty and clean homes, businesses, and streets, and men point out different risks and unclean elements, and discuss the consequences.

- Men identify ways they can improve cleanliness where they live (e.g. get a garbage can) and set and execute these plans as homework, to be followed up with home visits.

5. *Anger Management.* This module discusses the causes and effects of anger, the problems with acting out in ways they may later regret. It also provides participants with tools to manage their anger.

- Men discuss the signs and indications of anger, in themselves and others, through discussion and role playing. Facilitators show pictures of angry faces and situations, and men interpret them. The aim is to make them cognizant of these signs.

- Men discuss the causes of anger, and learn to link some of their actions to other people's anger.

- Men discuss and role play the negative consequences of aggression and violence, or share experiences from their own life.

- Men practice nonaggressive responses to angry confrontations in class, such as learning to distract or calm oneself (walking away, doing other activities, starting discussions and de-escalating, or practicing breathing techniques). Men practice these techniques as homework.

6. *Self-Esteem.* This module emphasizes the need for participants to discover themselves in order to begin the path to recovery. This module links their behavioral changes to respect, pride, and confidence.

   - The facilitators try to link poor self-image directly to many of the behaviors they have discouraged in previous modules, both as a cause and consequence.

   - Men discuss ways they can build self esteem, make plans, and execute them as homework.

   - Facilitators work with men to identify worthwhile skills and characteristics they hold that are worthy of others' respect.

   - Men practice shopping for goods in a supermarket or shop as one of the first exposure activities. They work through successes and failures as a group and try again, sometimes with the help of a facilitator.

7. *Planning.* Reviews the steps and components necessary for planning and implementation. The goal of this module is to build participants' capacity to develop short- and long-term plans and understand the processes involved in executing these plans.

   - Planning skills are commonly taught in CBT programs as a method to build new skills. At its most basic, this involves helping the men break down larger plans into smaller steps and helping them work through ways to accomplish those steps, positively reinforcing successes and helping them process challenges

and setbacks, often as a group. Men give examples and discuss them together. Another example: Small groups of men are tasked with organizing activities, such as a football match. The larger group listens to the different plans and critiques them.

- As homework assignments, initially men are tasked with simple tasks (create a short term survival plan for feeding yourself or your family), and then more complex tasks (such as a business plan or home garden).

- Men are also tasked with identifying a successful friend or family member and determining what steps led to their success. A motivational speaker (usually a past graduate) is also invited to talk about the steps involved in their success and their learnings and setbacks.

8. *Goal Setting.* The module outlines tools participants can use to develop goals, objectives, and indicators for measuring success in their own lives.

- Participants are taught what short and long term goals are (through discussion and examples) and how to set reasonable short- and long-term goals (such as feeding their family, or starting a garden).

- First participants practice setting goals and making plans, and then the larger group discusses and critiques them. Participants then set their own small, short term goals (e.g. changing a behavior, reconciling with a family member, or saving a certain amount this week) and execute these as homework, processing successes and failures as a group.

- Participants discuss the characteristics of good goals (e.g. achievable, measurable, time-bound) and revise goals and plans. They are given poor goals as a group and practice turning them into better goals. Another motivational speaker is used to discuss the role of goal setting in their own life.

9. *Money Business.* Stresses the importance of engaging in positive spending habits and appropriately managing money. Impulsive spending habits are emphasized. Participants are taught to make plans and prioritize their needs and wants prior to spending their money.

- Men engage in exercises to track their own recent spending to see where their money has gone. They discuss the use and misuse of their own money. As

a group they discuss regrets and bad decisions and work through the negative consequences. These are illustrated dramatically through role-playing and skits, followed by discussion.

- Later discussion, role playing and skits focuses on techniques for resisting peer pressure and temptation. There is also testimony from a motivational speaker, usually a past graduate of the program.

10. *Money Saving.* The module introduces participants to various saving options and encourages them to reflect on the most suitable saving method for their lives. They practice interactions in informal and formal financial institutions.

- Men discuss the reasons for and advantages of saving and it is explicitly linked to positive self image and esteem in the community. There is another motivational speaker.

- Men learn techniques for saving safely at home without formal institutions. They learn to set and execute saving plans, using their goal setting and planning skills.

- Homework assignments involve saving money they would have otherwise used on things they regret (identified in the previous module). Homework also involves trips to the bank and informal lenders. Prior to these assignments they meet and role play in groups, and their strategies are discussed and critiqued by the larger group. There is also a focus on appropriate presentation and image in these outings.

11. *Challenges and Setbacks.* The module explores potential challenges and setbacks they will face and has them practice positive coping mechanisms needed to effectively overcome them. Challenges and setbacks are framed as a test of one's maturity, potential, and abilities, and an opportunity for improvement.

B.3   Cash grants

We contracted the international non-profit Global Communities (GC) to conduct the registration and cash distribution, as well as oversee NEPI's financial management and implementation schedule. We did so for several reasons:

1. To keep the therapy and the research teams distinct from cash distribution;

2. To coordinate registration and implementation of the two activities;

3. To relieve the research team of project and financial management of the interventions; and

4. To make the intervention as close as possible to a real-world, replicable intervention by other non-profit or state organizations.

For safety, GC developed a highly structured system of cash distribution. GC staff held cash in a car that moved around the neighborhood, to avoid theft. A lottery team with the men gave grant winners a voucher, and put them on a motorbike taxi that was then directed to the street corner where the car with the cash awaited. They were told to approach the car (which had an identifying mark such as a red bag on the dash), hand over their voucher, and receive their cash. The car would then move to a new corner, whose location would be relayed by mobile phone, and the process would repeat.

Anyone who was assigned to the cash treatment but was not present on the day of disbursal was still eligible for the grant. GC attempted to locate them for up to three weeks afterward, and generally succeeded.

C   Formal theoretical model

C.1   Setup

We model an individual's choice between legitimate business and illicit activities under different conditions—with and without time inconsistency, and with and without financial market imperfections—and assess the predictions for a number of common labor market and crime-reducing interventions: greater punishment, increasing productivity in legitimate business (e.g. through technology or skills improvement), cash or capital transfers, and interventions that shape preferences—either time preferences or personal preferences against illegal behavior.

We use L^b and L^c to denote time spent in legitimate activities (such as petty business) and illegitimate activities (such as crime). Legitimate business produces revenue according to production function $F(\theta, L_t^b, K_t)$, where θ is productivity or individual ability and K is accumulated capital used in business. A person's decision to participate in illegal activity is motivated by the potential gains and costs from such activity. Gains include the expected illegitimate payoff per hour spent in illegal activities, w. Costs include the possibility of apprehension and conviction, which occurs with probability, ρ, and implies a

penalty, fL^c_{t-1}. Thus the penalty for criminal behavior is a linear function of hours spent in criminal activities in the previous period[34] The individual's total expected earnings from legitimate and illegitimate activities are $y_t \equiv F(\theta, L^b_t, K_t) + w_t L^c_t - \rho f L^c_{t-1}$. . In addition to investing in business, the individual can also invest or borrow through a riskless asset with constant returns $1 + r$. At each period t, the individual decides how much to invest for next period a_{t+1} and reaps interests ra_t from last period's investments.

Individuals have utility function $U(c, l, \sigma L^c)$, where c denotes consumption and l denotes time for leisure. We also allow for individuals to have direct disutility from engaging in crime, as measured by σL^c, where $\sigma > 0$ implies that implies that illicit work induces some internal penalty such as shame, though in principle it could also reflect social penalties such as a loss of esteem or exclusion from peers and other social networks. We make the standard assumption that $U'_c \geq 0$, $U'_l \geq 0$, $U'_{\sigma L^c} \leq 0$, $U''_{cc} < 0$, $U''_{ll} < 0$, $\partial^2 U / \partial L^2_c \leq 0$ and $F'_\theta \geq 0$, $F'_L \geq 0$, $F'_K \geq 0$, $F''_{\theta\theta} < 0$, $F''_{LL} < 0$, $F''_{KK} < 0$, and $F''_{\theta L} \geq 0$, $F''_{\theta K} \geq 0$, $F''_{LK} \geq 0$.[35] We allow for the individual to have quasi-hyperbolic (β, δ) preferences.

We first consider the case without any uncertainty. The individual's problem is:

$$\max_{c_t > 0, 0 \leq l_t \leq \bar{L}, L^b_t, L^c_t, K_{t+1}, a_{t+1}} \quad U(c_t, l_t, \sigma L^c_t) + \beta \sum_{i=1}^{\infty} \delta^i U(c_{t+i}, l_{t+i}, \sigma L^c_{t+i})$$

$$s.t. \quad c_t + a_{t+1} + K_{t+1} = F(\theta, L^b_t, K_t) + w_t L^c_t - \rho f L^c_{t-1} + (1+r)a_t \quad for\ each\ t$$

$$a_0 \quad given$$

where $L^b_t + L^c_t + l_t \equiv \bar{L}$.

---

[34]One reason for this modeling choice is because we want to explore the role that quasi-hyperbolic preferences play in the decision to commit crimes when the punishment is in the future not the present.

[35]For ease of analysis, we also assume that the marginal return to capital is infinity for the first unit of capital invested in business, and that as long as there is positive capital input, marginal product of labor for the first unit of labor will be infinity, i.e. $\lim_{K \downarrow 0} F'_K(\theta, L^b, K) = +\infty$ for all L^b and $\lim_{L^b \downarrow 0} F'_K(\theta, L^b, K) = +\infty$ as long as $K > 0$. This assumption guarantees that investments and hours in business will always be positive.

C.2 Occupational choice (and interventions) among time consistent individuals

Without credit constraints

Without time inconsistency ($\beta = 1$) or credit constraints, the set of optimality conditions are:

$$\frac{U_l'(t)}{U_c'(t)} = F_{L^b}'(t) \qquad if\ L_t^b > 0 \tag{1}$$

$$\frac{U_l'(t)}{U_c'(t)} - \sigma\frac{U_{\sigma L^m}'(t)}{U_c'(t)} = w_t - \frac{\rho f}{1+r} \qquad if\ L_t^m > 0 \tag{2}$$

$$1+r = F_K'(t+1) \qquad if\ K_{t+1} > 0 \tag{3}$$

$$\frac{U_c'(t)}{U_c'(t+1)} = \delta(1+r) \tag{4}$$

$$c_t + a_{t+1} + K_{t+1} = F(\theta, L_t^b, K_t) + w_t L_t^c - \rho f L_{t-1}^c + (1+r)a_t \tag{5}$$

where for ease of notation, we use $U(t)$ to denote $U(c_t, l_t, \sigma L_t^c)$ and $F(t)$ to denote $F(\theta, L_t^b, K_t)$. Since we modeled crime punishment as a potential reduction in future wages, the risk neutral individual will view crime as an occupation with a discounted wage $w_t - \frac{\rho f}{1+r}$.

To find the marginal conditions for engaging in each sector, we first consider the case where illicit activity is not feasible. This would arise naturally if the probability of apprehension is high enough and punishment is heavy enough that $w \ll \frac{\rho f}{1+r}$. In this case the decision to engage in business depends on productivity θ, wealth level and the returns on other financial assets r. We use c^{ba}, L^{ba} and K^{ba} to denote consumption, labor and capital level in this scenario. Each period t, the individual chooses L_t^{ba} to satisfy $\frac{U_l'(c_t^{ba}, \bar{L} - L_t^{ba}, 0)}{U_c'(c_t^{ba}, \bar{L} - L_t^{ba}, 0)} = F_{L^b}'(\theta, L_t^{ba}, K_t^{ba})$ taking K_t^{ba} as given, and he chooses capital investment K_{t+1}^{ba} to satisfy $F_K'(\theta, L_{t+1}^{ba}, K_t^{ba}) = 1 + r$, taking expected L_{t+1}^{ba} as given.

Now, taking levels of c^{ba}, L^{ba} and K^{ba} as given, we then look at individuals' decision to engage in crime. Individuals will engage in illicit activities if and only if:

$$w_t - \frac{\rho f}{1+r} \geq \frac{U_l'(c_t^{ba}, \bar{L} - L_t^{ba}, 0)}{U_c'(c_t^{ba}, \bar{L} - L_t^{ba}, 0)} + \sigma\frac{-U_{\sigma L^m}'(c_t^{ba}, \bar{L} - L_t^{ba}, 0)}{U_c'(c_t^{ba}, \bar{L} - L_t^{ba}, 0)} \tag{6}$$

which says expected returns from crime are higher than the highest possible marginal rate of substitution between leisure and consumption the individual can achieve without engaging in crime. Since $-U_{\sigma L^m}'/U_c' > 0$, a rise in σ means more people will drop out of

crime.

If condition (6) is satisfied and if $K_t > 0$, the individual then chooses L_t^b and L_t^c such that the marginal product of labor in business equals his expected marginal gains from crime, which also equals his marginal rate of substitution between leisure and consumption: i.e. conditions (1) and (2) will be satisfied. Notice L_t^c may not always be positive. The individual will not engage in crime if any or all three of the following happens: w_t is very low relative to the probability of apprehension ρ and punishment f; productivity in business θ is very high; the degree of aversion to crime σ is very high.

Capital investment and hours in business will satisfy condition (3). Notice that w, ρ and f will not affect returns to investment in business.

Interventions that increases the disutility of crime or the size or probability of punishment will reduce time devoted to in crime, but will have no effects on returns in business.[36] However, interventions that increase business productivity θ will not only induce more investment in business, but also reduce involvement in crime. In other words, $\frac{\partial L^c}{\partial \sigma} < 0$, $\frac{\partial L^b}{\partial \sigma}$ is ambiguous, $\frac{\partial L^c}{\partial \theta} < 0$ and $\frac{\partial L^b}{\partial \theta} > 0$. Finally, interventions that provide capital or liquid financial assets, such as a cash windfall, will not affect occupational choice at all, since the individual will already be working at his optimal level in both sectors. The windfall will simply be consumed and saved.

With credit constraints

In this section we consider the model with a simple credit constraint in the form of $a_t \geq 0$–individuals are unable to borrow in any period. We focus our attention on individuals whose initial a_0 is low enough that at some point in his life, the credit constraint is binding. Credit constraints will affect optimal conditions (2) and (3). The optimal condition for capital investment (3) becomes

$$F'_K(\theta, L_t^b, K_t) = max\{1 + r, \frac{1}{\delta}\} \qquad if \ K_{t+1} > 0$$

and the optimal condition for hours in crime (2) becomes

$$\frac{U'_l(t)}{U'_c(t)} - \sigma \frac{U'_{\sigma L^m}(t)}{U'_c(t)} = w_t - \frac{\rho f}{max\{1 + r, \frac{1}{\delta}\}} \qquad if \ L_t^m > 0$$

Notice that $max\{1 + r, \frac{1}{\delta}\} \geq 1 + r$ and $w_t - \frac{\rho f}{max\{1+r,\frac{1}{\delta}\}} \geq w_t - \frac{\rho f}{1+r}$. For the impatient

---

[36]The level of investment in business may change depending on the shape of the utility and production functions, but the returns to investment will not change.

individuals whose $\frac{1}{\delta} > 1+r$, their optimal level of capital investment will be lower than the baseline case because of the credit constraint. They are also have a higher expected returns from crime than in the baseline case, because the low level of business investment also forces them to put a higher discount rate on potential future punishment from crime.

Critical condition (6) becomes

$$w_t - \frac{\rho f}{max\{1+r, \frac{1}{\delta}\}} \geq \frac{U_l'(c_t^{ba}, \bar{L} - L_t^{ba}, 0)}{U_c'(c_t^{ba}, \bar{L} - L_t^{ba}, 0)} + \sigma \frac{-U_{\sigma L^m}'(c_t^{ba}, \bar{L} - L_t^{ba}, 0)}{U_c'(c_t^{ba}, \bar{L} - L_t^{ba}, 0)}$$

Credit constraints induce more individuals who would otherwise not engage in crime to commit crime. For the impatient individuals, credit constraints increase their hours in crime and reduce their capital investments and hours in business activities.

Interventions that ease the credit constraint, including cash windfalls, will induce more investment in business and reduce involvement in crime. As in the baseline case, $\frac{\partial L^c}{\partial \sigma} < 0$, $\frac{\partial L^b}{\partial \sigma}$ is ambiguous, $\frac{\partial L^c}{\partial \theta} < 0$ and $\frac{\partial L^b}{\partial \theta} > 0$; however, the magnitude the effects of a change in σ or θ will be greater than in the baseline case; the magnitudes also increases with the degree of impatience: $\frac{|\partial L^c/\partial \sigma|}{d\delta} < 0$, $\frac{|\partial L^c/\partial \theta|}{d\delta} < 0$ and $\frac{|\partial L^b/\partial \theta|}{d\delta} < 0$ (notice that the lower the value of σ, the more impatient the individual).

C.3 Occupational choice (and the effects of interventions) under time inconsistency

Without credit constraints

Time-inconsistent individuals ($\beta < 1$) will be more reckless in the present. Intuitively, the smaller is β, the more individuals want to enjoy higher consumption today at the expense of future consumption, which means they will borrow more, save less, invest less in business and/or involve more in criminal activities. However, as long as there is a perfect financial market, no one will change their business or criminal activities in order to consume more today–they will simply borrow more (or save less) today through the financial market.

In terms of optimal conditions, in the absence of any credit constraint, the only condition that changes is equation (4), which becomes

$$\frac{U_c'(c_t, l_t, \sigma L_t^c)}{U_c'(c_{t+1}^P, l_{t+1}^P, \sigma L_{t+1}^c)} = [\frac{\partial c_{t+1}}{\partial W_{t+1}} \beta \delta + (1 - \frac{\partial c_{t+1}}{\partial W_{t+1}}) \delta] \cdot (1+r)$$

where W_t denotes total wealth at time t, c_{t+1}^P denotes the individual's predicted future

decision about c_{t+1} at time t. For the sophisticates $c_{t+1}^P = c_{t+1}$ while for the naifs $c_{t+1}^P > c_{t+1}$. Compared with the baseline case, the discount factor δ is replaced by the effective discount factor $\frac{\partial c_{t+1}}{\partial W_{t+1}}\beta\delta + (1 - \frac{\partial c_{t+1}}{\partial W_{t+1}})\delta$, a weighted average of the short-run and long-run discount factors $\beta\delta$ and δ where the weights are the next period marginal propensity to consume out of total wealth.

Notice that neither condition (2) nor condition (3) changes, as long as we have no credit constraints. Compared with the baseline, time inconsistency alone will not affect criminal activities or business investment. It would only change the level of savings or debts.

In this case, interventions that aim to correct time consistency will have no effects on either business investment or criminal activities, but will have an effect on consumption, savings and income.

With credit constraints

With credit constraints, in addition to equation (4), optimal conditions (2) and (3) will change as well. Let $\Delta = \frac{\partial c_{t+1}}{\partial W_{t+1}}\beta\delta + (1 - \frac{\partial c_{t+1}}{\partial W_{t+1}})\delta$ be the effective discount factor under (β, δ) preferences and $\tau = \frac{U_c'(c_t, l_t, \sigma L_t^c)}{U_c'(c_{t+1}^P, l_{t+1}^P, \sigma L_t^c)} \cdot \frac{1}{\Delta}$, where c_{t+1}^P denotes the individual's predicted future decision about c_{t+1} at time t. With credit constraints, the Euler equation (4) becomes

$$\tau \geq 1 + r \qquad with\ equality\ iff\ a_{t+1} > 0$$

and conditions (2) and (3) become

$$\frac{U_l'(t)}{U_c'(t)} - \sigma\frac{U_{\sigma L^m}'(t)}{U_c'(t)} = w_t - \frac{\rho f}{\tau} \qquad if\ L_t^m > 0$$

and

$$F_K'(\theta, L_t^b, K_t) = \tau \qquad if\ K_{t+1} > 0$$

In addition, critical condition (6) will change accordingly, with $1 + r$ replaced by τ.

Compared with the baseline case, $\tau > 1 + r$ as long as an individual is credit constrained (i.e. has no savings). The level of τ will be higher for the sophisticates than for the naifs. However, regardless of their level of sophistication (i.e. the way individuals set their expectations for their future behavior), we know for sure that $\tau > \frac{1}{\delta}$, and the smaller β is (i.e. the more time inconsistent), the higher τ will be.

Compared to the time-consistent credit constrained case, fewer individuals will invest in business, more individuals will engage in crime, business investment levels will be lower,

and hours in crime will be higher for everyone. The difference increases with the level of inconsistency (i.e. decreases with β).

Interventions that improve time consistency will shift people away from crime towards business. So will increasing the disutility of crime (though, as in the case without time inconsistency, while $\frac{\partial L^c}{\partial \sigma} < 0$, $\frac{\partial L^b}{\partial \sigma}$ is ambiguous). Increasing business productivity will have similar effects as before: $\frac{\partial L^c}{\partial \theta} < 0$ and $\frac{\partial L^b}{\partial \theta} > 0$. In all of these cases, however, the magnitudes the effects of a change in σ or θ will be greater than under time consistency, and the magnitudes also increase with the both degree of impatience and the degree of time inconsistency: $\frac{|\partial L^c/\partial \sigma|}{d\beta} < 0$, $\frac{|\partial L^c/\partial \theta|}{d\beta} < 0$, $\frac{|\partial L^b/\partial \theta|}{d\beta} < 0$, $\frac{|\partial L^c/\partial \sigma|}{d\delta} < 0$, $\frac{|\partial L^c/\partial \theta|}{d\delta} < 0$ and $\frac{|\partial L^b/\partial \theta|}{d\delta} < 0$. Notice that the lower the value of β, the more time inconsistent the individual is, and similarly, the lower the value of σ, the more impatient the individual is.

C.4   Introducing uncertainty and risk aversion

Three potential sources of risk are uncertainties in business productivity θ, wages from criminal activities w, and the potential punishment after apprehension f. We assume that decisions on business investment and hours in both sectors are made before risks are realized, and that θ, w and f follow independent stochastic processes.

With uncertainties in both the business and illicit sector, business investment and hours in both sectors depend on the variance of returns in both sectors and the level of initial wealth a_0. If both sectors are sufficiently risky, then those with high levels of wealth a_0 will turn away from both activities by reducing K, L^b and L^c and investing instead in other riskless assets. K, L^b and L^c will all be lower than the cases without risk. Those with low levels of initial wealth will not be able to live off savings alone, so they will have to invest more in either or both sectors, depending on the relative riskiness of the two sectors. As long as both sectors are similarly risky, K, L^b and L^c will all be higher; otherwise, if one of the sectors is less risky than the other, individuals will invest more time in that sector. $\frac{L^c}{L^b+L^c}$ will be lower than in the case without uncertainty if returns to crime are more volatile than business returns. One special case would be if individuals face a significantly positive chance of death after committing any crime. This is the equivalent of saying $f = +\infty$ with strictly positive chances. In this case hours in crime will be reduced to zero as long as the probability of apprehension is positive, $\rho > 0$.

With the presence of risk, inventions in θ will have greater effects, because an increase in θ now also makes business *relatively* less risky. A rise in σ will also have a bigger effect than without uncertainty, because risk aversion will reinforce the rise in aversion and further reduce hours in crime.

D Measurement

D.1 Self-reported non-cognitive skills

Because all personality questions were selected from questionnaires used in the United States, they were first translated into Liberian English by the enumerators, and then pre-tested with individuals selected from the same population as the youth in our study. To ensure that questions continued to assess the original underlying constructs two checks were performed. First, within the pre-test data we ensured that groups of questions were correlated or anti-correlated as one would expect given the underlying personality measure (e.g., impulsivity was negatively correlated with conscientiousness). Second, in Phase 2 we performed factor analyses to ensure that within scales, questions were answered similarly. Internal consistency within scales was confirmed in this way.

D.2 Cognitive abilities

The behavioral protocol consisted of several interactive "activities", some drawn from psychology and some from economics.[37]

Spatial comprehension and general problem solving. We asked subjects to complete a simple 6-piece jigsaw puzzle, where the image was a photograph of a familiar local setting such as a filling station. For individuals familiar with puzzles, this would be a simple test, however, none of the subjects had encountered jigsaw puzzles before, and thus this served as a test of problem solving within the novel testing environment. Subjects received a score (1 or 0) indicating whether the puzzle was correct, and a time to completion. If the subject was not finished when 5 minutes had passed they received a score of 0 (incorrect) and their time to completion was 5 minutes. On average subjects took 2.05 minutes (SD = 1.14 min) to complete the puzzle. Average accuracy on the puzzle was 90%.

---

[37]Across all behavioral tests administration was standardized. First, a clinical psychologist and economist trained enumerators in test administration. Next, in collaboration with experienced enumerators and research assistants, a comprehensive protocol was developed and used by all future enumerators. Enumerators were also instructed to answer clarifying questions and were taught the over-arching concept within each game so they could address questions/alleviate concerns without straying from the central concepts of the tests. This tight control over the testing situation allowed us to collect relatively sophisticated measures of cognitive function and behavioral responses to rewards in a constrained and otherwise under-resourced testing environment.

Planning behaviors. We used a series of mazes to test planning behavior. Again, mazes were unknown to nearly all respondents. Similarly to the puzzle, subjects were shown an example maze on paper and then given 2, 2, and 3 minutes respectively to complete increasingly difficult mazes. Each had two entry points, one of which almost immediately led to a dead end. The main outcome of the mazes was the subject's ability to pause and plan their approach before completing the maze (i.e. did they plan their approach before choosing a starting point). As outcomes, we measure "time to first touch", or the amount of time spent planning prior to engaging in the maze; and number of mistakes (or "backtracks") in Maze 3, the hardest maze, which required the most planning and by which time participants had learned the concept of the maze. On average subjects took 18 seconds to plan for Maze 3 (SD = 23 seconds).

Behavioral inhibition and cognitive flexibility We developed the "arrows game", a modified directional Stroop task, a class of tasks that assess inhibitory control. Here subjects were shown a sequence of large black or white arrows that pointed either up or down and were first told to respond "up" or "down" to each arrow ("arrows baseline"). In the second version they were again shown the arrows but now were told to state the opposite direction; this constitutes producing the less common response while suppressing the more common response and is an assessment of inhibition ("arrows inhibition"). Finally, in a third version subjects were told to switch between two approaches: if the arrow was white they were to state the actual direction, but the opposite direction if the arrow was black. This is commonly called 'switching' and is an assessment of cognitive flexibility, the ability to move rapidly between two goals as the situation demands ("arrows switching"). For each version, the outcome data included total time to completion and the number of correct/incorrect responses out of 32 arrows. On average subjects made .33 errors (SD = 1.5) on arrows baseline, 2.4 errors (SD = 3.5) on arrows inhibition, and 3.9 errors (SD = 3.9) on arrows switching. Arrows took on average 25 seconds (SD = 17.7), 38 seconds (SD = 45.8), and 46 seconds (SD = 28.7) for baseline, inhibition, and switching separately.

Working memory. Working memory is the ability to hold something in mind when it is no longer present in the environment and then manipulate it. The digit span task is an assessment of working memory. The digit span tasks involved the enumerator saying a random sequence of digits (1-9) out loud with a short pause between each digit, followed by the respondent repeating them back either in the same (forward-digits) or the reverse (backwards-digits) order. The enumerator began by giving two 2-digit numbers (one at a

time) and recording the responses. If the subject correctly reported either of the numbers back, the enumerator would do the same with 3-digit numbers, and so on up to a maximum of 9 digits. As soon as the subject incorrectly reported both examples at a given level or span the enumerator moved on to the next activity (backwards-digits). The reverse digit span was done the same way, except that the subject was instructed to repeat the digits in the opposite order that the enumerator gave them (e.g., "three, zero, one") On average subjects were able to remember 5.5 digits forward (SD = 1.23) and 3.33 digits backwards (SD = 1.03).. Each activity existed as two slight variants (e.g. changing the numbers in the gambles). These activities were alternated in the 2 versus 5-week endlines and the 12 versus 13-month endlines, so that participants were never asked identical questions too close together in time.

D.3 Economic preferences

The behavioral protocol also measure time and risk preferences.

Intertemporal choices. The penultimate activities were incentivized inter-temporal choices and hypothetical large-magnitude inter-temporal choices. These consisted of a series of binary choices between money at one point in time versus more money later in time.

Subjects were told that one of the questions across the next few activities would be picked for payout, with their choice implemented, so that they should pay careful attention to their decisions. We had told subjects that if one of the inter-temporal tasks was chosen for payout, and if their individual choice implicated a delayed reward, that we would come back and find them at the appointed time, in their own environment, to pay them.[38]

In this activity, after a clarifying example, subjects were asked four questions: first a choice between money now and more money in two weeks; second between two weeks and four weeks; and finally one more question for each of these pairs of delays, but with the numbers modified depending on their first answer (i.e. if they chose to wait, then they were asked again but with a lower reward in the future). This bifurcating design allowed

---

[38]In fact, for logistical reasons, we always chose one of the risk questions for actual payout. Although we did not technically lie at any point, this could be construed as minor deception. None of the respondents brought this up, even after having gone through the process five times. By the endline stage, they knew us fairly well and knew that we were able to track them (and that we had paid them everything we had promised them in the past). We also directly asked them, at the end of the behavioral protocol, whether they believed that we would return to find and pay them, if necessary. In addition we asked whether they expected the program to deliver any future benefits, in case such an expectation might also have influenced their stated choices.

us to glean as much information as possible about their preferences with as few questions as possible, and we pretested the numbers in order to maximize the variance in responses.

The enumerators then told respondents that they were done with the "real money" games but that they would be asked a series of hypothetical inter-temporal choices for larger amounts of money (on the order of US$10-30, about a week's wages). This was organized as two lists of 11 binary decisions, with a fixed amount right now versus a varying amount in two weeks (or two weeks versus four weeks for the second list). The delayed amount started as strictly less than the sooner amount, then equal to, and then larger and larger until it was four times as big.[39]

Risk preferences. The final activity was similar except that it examined risk preferences: subjects were shown a flashcard with an example of a binary lottery, along with an explanation of how it worked. We used the Liberian term "lucky ticket" to refer to this process. Then they were reminded that real money was (probabilistically speaking) at stake, and they were asked to choose between two binary lotteries, one with higher variance and higher expected value than the other. Depending on their choice, they were asked to make a second choice (again in a bifurcating design). Next, in order to assess loss aversion, everyone was asked to choose between doing nothing and playing a lottery in which they could either lose a small amount or gain a larger amount.

E   Survey data validation details

E.1   Variables

We selected six survey variables for validation, all with recall periods of two weeks. We attempted to choose a set of outcomes important to the analysis with varying degrees of salience (i.e. memorability) and potential social stigma and experimenter bias. The variables were:

1. *Stealing.* The corresponding survey questions asked, in the last two weeks, if the respondent stole someone's belongings when they were not paying attention ("corrected someone's mistake" in the Liberian English vernacular), shoplifted an item ("took something from behind someone that's not for you"), or deceived or conned

---

[39]In part this final activity was a way to quickly collect more data about a key outcome variable, and in particular to do so with substantial (hypothetical) magnitudes at stake. However in part it was a robustness check, since it is notoriously difficult to interpret the results of real-money inter-temporal choices, given that there is almost always some element of trust and risk involved as well.

someone of money ("cheated or scraped from people").[40] Based on our fieldwork, we hypothesized that stealing would be the most salient and least socially desirable of all six measures.

2. *Gambling.* The corresponding survey question was, "In the last two weeks, how many days did you gamble, including betting like football games and other things?" Beforehand, we hypothesized gambling had a lower level of salience and sensitivity than stealing, but was still stigmatized somewhat.

3. *Marijuana use.* The corresponding survey questions were, "In the last 7 days how many times did you smoke opium?" and, "What about the 7 days before that?" Opium is the vernacular for marijuana in Liberia, and does not imply an actual opiate. Marijuana use is not socially acceptable across Liberian society overall, but is fairly prevalent in our target demographic. We initially hypothesized underreporting could arise not so much from social stigma but from the discouragement of drug use in the therapy treatment.

4. *Homelessness.* The corresponding survey question was, "In the last 2 weeks, have you ever had to sleep outside, or on the street, or in a market stall because you had no other place to sleep or stay?" This is a salient variable where we hypothesized respondents might have under-reported from embarrassment or over-reported in order to appear more needy (and eligible for more programs).

5. *Phone charging.* The corresponding survey question was, in the context of an expenditure portion of the survey, "In the last 2 weeks, how many times did you charge phones?" In the vernacular, this corresponds to taking one's phone to a kiosk with electricity where one pays a small fee to recharge the battery. This is a common and routine expense for many Liberians, without stigma but possibly not very memorable. 38% of our sample had a mobile phone at the endline, and 38% reported charging a phone in the last two weeks.

6. *Video Club Attendance.* The corresponding survey question was, "In the last 2 weeks, how many times did you go to the video club?" These clubs are private businesses where one can go to watch a movie, television show, or football match for a small fee. This is a popular and socially acceptable pastime, as most Liberians do not

---

[40]The survey also measured more serious forms of theft, such as armed robbery, but our qualitative validation focussed on non-violent theft.

have electricity or home entertainment. Salience was unclear but likely greater than phone charging.

The program also intended to change political behavior, particularly participation in election violence, association with ex-military commanders, and participation in riots. These would normally be candidates for qualitative validation. In this particular instance, there were few opportunities for political violence during our study period, especially as the election turned out to be a peaceful affair. Also, our three-phase design meant that opportunities for political violence would have varied by phase. As such, political violence was not an ideal candidate for field testing the method in this instance.

E.2 Validator staff

Eight different people did the validation over the two years of data collection. Typically two were active during each validation round. All but one were men, and all had a high school or some post-secondary education. Two of the men completed roughly half the validations with the remainder doing roughly 10 to 20% each.

We selected validators from the study's best survey enumerators. We had hired and trained half of these to do more standard qualitative data collection also, and this prior qualitative research experience generally improved the quality of their work. To find these validators, we trained roughly two to three times the number of people needed from the pool of research staff, selecting only those with the most natural questioning and rapport-building skills for the validation exercise. Each received at least 10 days of training on the validation methods involving both classroom learning and extensive field training. Like any qualitative study, we believe staff recruitment and training to have been among the most important tasks and also the largest start-up cost of this method.

E.3 Approach

For each respondent, validators tried to determine whether the respondent had engaged in any of the measured behaviors, even once, in the two weeks preceding the respondent's survey date. We found it optimal for validators to visit each respondent four times, on four separate days, with each visit, or "hangout session" lasting approximately three hours. The validator aimed to begin hanging out the day after subjects completed their quantitative surveys and to conduct all four visits in the days following the respondent's endline survey date.

On the first visit validators would obtain verbal consent for hanging out and learning more about participants' lives. We designed the consent script to be informal, and explained that the goal of hanging out with the respondent was to talk about some of the same things they discussed in the survey, but also to get qualitative information about people's lives. In addition to this verbal consent, the more formal consent form delivered with the survey said that qualitative staff may come and visit them again to gather more information.

Validators deliberately avoided the feeling of a formal interview. Validators would typically shadow the respondents as they were going about their business, rather than sit down for a formal interview. Validators sometimes made notes during longer interviews, but only in isolated areas out of sight from the respondent, such as a toilet stall or teashop. If validators were unable to find a secluded area in which to take notes, they sometimes recorded information in their cell phones, pretending to send a text message.

The main approach was to engage in casual conversation on a wide range of topics, including the six target topics/measures. The target topics were raised mainly through indirect questions while informally chatting and conversing. For example, validators typically started conversations with discussions of family. This was both customary among peers in Liberia and a sign of interest in their lives and respect. It was also a stepping stone for discussing the target behaviors—either because the validator can discuss an issue in their family (someone engaging in one of the activities) or how the respondent's family feels about their current lifestyle and circumstances.

In general, validators found it helpful to tell respondents stories or scenarios about another person, or themselves, related to the target measures, then steer the conversation to get information about how respondents have behaved in similar situations, eventually discussing the past two weeks. Validators were careful to present these behaviors and incidents in a non-stigmatized light, for instance by discussing a friend who stole in order to get enough to eat, or how they themselves had periods of homelessness or used drugs and alcohol. Validators found these personal stories (all of which were truthful) and genuineness were essential to building rapport and trust.

Validators might hold these conversations once or twice over the three hours, spending perhaps twenty or thirty minutes in conversation each time, to avoid unnaturally long or awkward conversations. The validator spent the remainder of the three hours in the general vicinity, observing respondents engaging in their daily activities. This could involve taking a rest in the shade or in a tea shop (as is common) or engaging others in conversation. Validators would also try to talk casually with the respondent's friends, relatives,

or neighbors to learn about him (although we considered information from these second-hand sources as insufficient to support a conclusion about the respondents' behaviors, but merely as supporting information).

We found that building a rapport with participants in a short space of time was crucial to success. To develop trusting and open relationships, validators developed techniques, including becoming close to respected local community and street leaders, eating meals with participants, sharing personal information about themselves, befriending respondents' acquaintances, assisting subjects with their daily activities, and mirroring participants' appearances and vernacular, as appropriate. In addition, validators tried to maintain neutrality and openness while discussing potentially sensitive topics. For instance, conveying—through stories or otherwise—that illicit behaviors were not perceived negatively, allowed respondents to feel comfortable sharing their involvement in such activities. Validators did not lie to or deceive respondents, however.

E.4   Coding validated data

Validators were unaware of the respondents' survey responses, and formed their own opinions based on the information they collected about whether the respondent engaged in the six activities during the time period captured by the quantitative survey.

A core part of the validator training included logical reasoning, supporting their reasoning with evidence, and writing this down in a clear and structured manner. After each visit, validators made written notes about the relevant data collected, including evidence to support their conclusions, on a standardized form. At the conclusion of the four visits, the validator coded six indicators, one for each behavior, where "1" meant that he had relatively direct evidence that the respondent engaged in the behavior during the recall period, and "0" otherwise. This was a subjective judgment, but all submissions and coding decisions were reviewed with the authors case-by-case.

Over the course of the exercise, different measures offered different experiences and lessons. Because of its relative frequency and visibility, we suspect marijuana use was the easiest to directly observe. But validators found other behaviors straightforward to discuss in conversation. In the survey and (especially) the validation, phone battery charging led to the most confusion—in particular, did simply charging one's phone count, or did only paying to charge one's phone count? Paid charging was the focus of the survey question (it appeared in an expenditure survey module), but we were concerned that the validators would use a more expansive definition. We attempted to mitigate such differences through trainings and regular discussions on the coding.

Homelessness also proved somewhat challenging to measure and validate, as we discovered its definition is subjective. Circumstances arose that were somewhat ambiguous, such as having no home of one's own but regularly sleeping on a friend's floor or in an acquaintance's market stall. To account for the potential variability in perceptions of homelessness, validators were instructed to include as much information as possible about respondents' living situations in their summary reports. The authors then worked with validators to code a somewhat broad definition of homelessness that included any ambiguous circumstances. Prior to analysis, it was not clear whether survey respondents applied the same definition, and hence we err on the side of finding underreporting in the survey.

E.5   Validation sampling and non-response

In each endline survey round we randomly selected study respondents to be validated, stratified by treatment group.[41] In general, the validation sample was a balanced subsample of the full sample.

Table A.5 describes the samples selected for validation in each survey round over the course of the study. In total, we randomly selected 297 people for validation, and found 240 (81%).[42] We could not find 15 for even the endline survey. We could not validate a further 42 because they were difficult to find even immediately after the survey or (more commonly) because they lived a long distance away. In general, we surveyed respondents who had moved far out of Monrovia, but we were unlikely to validate them because of the time and expense and opportunity cost.

This attrition is an identification concern, but there is little evidence of biased attrition. Just as overall endline survey attrition was low and relatively uncorrelated with baseline characteristics, the excess validation attrition (those who were surveyed but not validated) was not robustly associated with baseline characteristics.

---

[41]We blocked respondents by survey window, and within each block we randomly selected validation subjects using a computer-generated uniform random variable. The selection was performed without replacement in a given pair of survey rounds (e.g. the short-term endline surveys in a given phase), but sampling was performed with replacement across survey rounds. Twenty subjects were validated in more than one round.

[42]We decided to select 297 based on a combination of logistical capacity and budget, alongside power calculations based on the earliest rounds of data collected.

Table A.5: Validation sample, totals and attrition

Phase	Round	Surveys Target #	Validation Selected	Validation Validated	Unfound at endline	Unfound for validation	% validated (all)	% validated (treatment)	% validated (control)
	3-week	100	0						
	5-month	100	24	18	0	6	75%	75%	75%
1	7-month	100	24	12	0	12	50%	50%	50%
	12-month	100	10	6	3	1	60%	63%	50%
	13-month	100	10	8	2	0	80%	86%	67%
	3-week	398	26	24	0	2	92%	94%	89%
2	5-week	398	27	17	0	10	63%	68%	40%
	12-month	398	28	25	2	1	89%	86%	100%
	13-month	398	44	38	1	5	86%	85%	91%
	3-week	501	0						
3	5-week	501	0						
	12-month	501	35	31	2	2	89%	89%	88%
	13-month	501	69	61	5	3	88%	88%	88%
All		4096	297	240	15	42	81%	81%	80%

Notes: The proportion selected in each round was principally a function of logistical feasibility (e.g. number of available staff), and in some none were selected. As procedures became more familiar and staff more experienced, more could be done over time. The percentage validated in the treatment group includes any treatment (cash, CBT, or both).

E.6 Limitations of the approach

While our qualitative assessment is that this validation exercise gave enough time to gather detailed, accurate information and fostered trust and frankness, there are nonetheless limitations to this approach.

1. The interviews may be intrusive and might disrupt respondents' daily activities, thereby altering the findings. To mitigate this risk, validators wore clothes that would blend in with their respondent's environment, and typically accompanied and assisted respondents in their activities as appropriate (e.g. helping a scrap metal collector scavenge).

2. The survey and validation questions might have been interpreted differently, making it difficult to compare results. As discussed above, phone charging and homelessness proved somewhat difficult to measure consistently. We used close consultations and reviews of the data, and focus groups with survey and validation staff, to maximize consistency.

3. The validation period came after the survey recall period and validators or respondents could have made errors about the relevant window of time (e.g. homelessness could have been observed the week after the survey, and inferred to the time of the survey incorrectly). This is most likely a source of random measurement error.

4. In principle the participant observation method, by building a rapport, could lead to a different source of measurement error by (for example) increasing social desirability bias. Our strong sense is that the opposite is true, that trust and rapport reduced the bias, but this is a subjective interpretation and not independently verifiable.

5. The validator staff could easily learn the subject's treatment status, simply through conversation, and in principle this could bias the validation data because of expectations.

6. The method is reliant on skills, persistence, luck, and best judgment of the validators and authors.

7. Like any qualitative work, this is not an off-the-shelf tool. To select and refine the variables, recruit and train validators, and monitor quality of the data requires the researcher to have some familiarity with the context and population and at least basic experience in qualitative data collection.

F Additional treatment effects analysis

F.1 The interaction of cash and therapy treatments

F.2 Robustness of treatment effects

Table D.1: Program impacts on personality traits, non-cognitive skills, and time and risk preferences

Outcome, z-score	Round	Control mean	Assigned to therapy (any)		Assigned to cash (any)		Both	
			ITT	Std. Err.	ITT	Std. Err.	ITT	Std. Err.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Index of traits associated with self control	2 & 5 wk.	-0.046	0.087	[.093]	-0.154	[.099]	0.103	[.136]
	12 & 13 mo.	-0.076	0.153	[.082]*	-0.025	[.086]	0.109	[.114]
Impulsiveness (9)	2 & 5 wk.	-0.042	-0.011	[.096]	0.185	[.102]*	-0.067	[.141]
	12 & 13 mo.	0.115	-0.175	[.087]**	0.007	[.089]	-0.040	[.118]
Conscientiousness (8)	2 & 5 wk.	-0.112	0.108	[.096]	0.046	[.097]	0.008	[.138]
	12 & 13 mo.	0.042	-0.068	[.092]	-0.032	[.094]	0.143	[.123]
Perseverance (7)	2 & 5 wk.	0.016	0.028	[.093]	-0.133	[.100]	0.149	[.141]
	12 & 13 mo.	-0.077	0.114	[.090]	0.056	[.090]	-0.067	[.124]
Reward responsiveness (8)	2 & 5 wk.	0.058	-0.073	[.101]	0.109	[.102]	-0.023	[.139]
	12 & 13 mo.	0.038	-0.160	[.091]*	0.079	[.089]	-0.154	[.119]
Neuroticism (8)	2 & 5 wk.	0.080	0.012	[.094]	0.022	[.097]	-0.114	[.136]
	12 & 13 mo.	-0.051	0.044	[.087]	0.037	[.091]	-0.227	[.121]*
Locus of control (8)	2 & 5 wk.	0.025	0.000	[.095]	0.059	[.100]	-0.153	[.142]
	12 & 13 mo.	-0.019	-0.031	[.088]	-0.106	[.086]	0.115	[.122]
Self esteem (8)	2 & 5 wk.	-0.054	0.110	[.090]	-0.012	[.091]	0.090	[.125]
	12 & 13 mo.	-0.113	0.078	[.090]	0.058	[.092]	0.052	[.121]
Patience in game play	2 & 5 wk.	-0.194	0.174	[.075]**	0.151	[.075]**	-0.082	[.103]
	12 & 13 mo.	-0.073	0.026	[.081]	0.134	[.077]*	-0.063	[.104]
Time inconsistency in game play	2 & 5 wk.	0.085	-0.042	[.079]	-0.118	[.082]	0.078	[.108]
	12 & 13 mo.	-0.018	-0.034	[.068]	-0.016	[.071]	-0.001	[.096]
Risk aversion in game play	2 & 5 wk.	0.147	-0.137	[.080]*	-0.110	[.084]	0.106	[.122]
	12 & 13 mo.	-0.033	-0.061	[.073]	-0.139	[.074]*	0.141	[.101]

Notes: Robust standard errors in brackets, clustered by individual. *** p<0.01, ** p<0.05, * p<0.1

Table D.2: Program impacts on anti-social behaviors

Outcome (No. of questions in indexes in brackets)	Round	Control mean	Basic ITT regression					
			Assigned to therapy (any)		Assigned to cash (any)		Both	
			ITT	Std. Err.	ITT	Std. Err.	ITT	Std. Err.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Index of all antisocial behaviors	2 & 5 wk.	0.171	-0.197	[.074]***	-0.089	[.073]	0.030	[.096]
	12 & 13 mo.	0.034	-0.091	[.081]	0.090	[.082]	-0.221	[.105]**
Aggressive and hostile behaviors (19), z-score	2 & 5 wk.	0.091	-0.173	[.069]**	0.009	[.071]	-0.041	[.095]
	12 & 13 mo.	0.188	-0.154	[.102]	-0.045	[.099]	-0.140	[.130]
Disputes/fights, past 2 weeks (6), z-score	2 & 5 wk.	0.366	0.089	[.106]	0.054	[.091]	-0.234	[.128]*
	12 & 13 mo.	0.273	-0.046	[.107]	0.084	[.104]	-0.160	[.133]
Carries weapon	2 & 5 wk.	0.157	-0.086	[.032]***	-0.044	[.034]	0.038	[.044]
	12 & 13 mo.	0.148	-0.058	[.029]**	0.044	[.032]	-0.051	[.042]
Usually sells drugs	2 & 5 wk.	0.170	-0.080	[.025]***	-0.040	[.027]	0.048	[.034]
	12 & 13 mo.	0.135	-0.037	[.027]	0.033	[.028]	-0.054	[.035]
# of thefts/robberies, past 2 weeks	2 & 5 wk.	2.580	-0.912	[.369]**	-0.842	[.382]**	0.540	[.474]
	12 & 13 mo.	1.876	0.000	[.362]	0.280	[.362]	-1.001	[.468]**
Arrested in past two weeks	2 & 5 wk.	0.139	-0.007	[.025]	0.007	[.025]	-0.010	[.033]
	12 & 13 mo.	0.115	-0.003	[.023]	0.009	[.023]	-0.036	[.031]

Notes: Robust standard errors in brackets, clustered by individual. *** p<0.01, ** p<0.05, * p<0.1

Table D.4: Program impacts on Attitudes and Drug Use

Outcome (No. of questions in indexes in brackets)	Round	Control mean	Basic ITT regression				Coefficient on a cash + therapy	
			Assigned to therapy (any)		Assigned to cash (any)			
			ITT	Std. Err.	ITT	Std. Err.	ITT	Std. Err.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Index of violent/criminal attitudes, z-score	2 & 5 wk.	0.172	-0.228	[.093]**	-0.225	[.094]**	0.254	[.127]**
	12 & 13 mo.	0.045	-0.080	[.080]	0.013	[.079]	-0.109	[.099]
Attitudes toward use of violence (11), z-score	2 & 5 wk.	0.021	-0.141	[.097]	-0.201	[.098]**	0.285	[.138]**
	12 & 13 mo.	0.051	0.017	[.100]	0.080	[.101]	-0.143	[.135]
Attitudes toward criminality (12), z-score	2 & 5 wk.	0.139	-0.177	[.099]*	-0.154	[.103]	0.090	[.133]
	12 & 13 mo.	0.044	-0.063	[.094]	-0.041	[.092]	-0.138	[.117]
Attitudes on political violence (10), z-score	2 & 5 wk.	0.217	-0.250	[.110]**	-0.204	[.107]*	0.256	[.149]*
	12 & 13 mo.	-0.005	-0.104	[.083]	0.008	[.083]	-0.047	[.102]
Positive peer qualities (20), z-score	2 & 5 wk.	-0.160	0.207	[.084]**	0.014	[.088]	0.013	[.118]
	12 & 13 mo.	0.040	0.011	[.081]	-0.070	[.083]	0.076	[.114]
Usually drinks	2 & 5 wk.	0.673	-0.102	[.037]***	-0.018	[.037]	0.095	[.052]*
	12 & 13 mo.	0.763	-0.072	[.039]*	-0.052	[.039]	0.082	[.053]
Usually uses marijuana	2 & 5 wk.	0.500	-0.102	[.034]***	-0.002	[.033]	-0.029	[.049]
	12 & 13 mo.	0.495	-0.029	[.033]	0.014	[.032]	-0.032	[.045]
Usually takes hard drugs	2 & 5 wk.	0.209	-0.029	[.025]	-0.013	[.027]	0.033	[.035]
	12 & 13 mo.	0.200	-0.007	[.028]	0.075	[.030]**	-0.062	[.040]

Notes: Robust standard errors in brackets, clustered by individual. *** p<0.01, ** p<0.05, * p<0.1

Table D.6: Program impacts on investment, income, and employment

Outcome (No. of questions in indexes in brackets)	Round	Control mean	Basic ITT regression					
			Assigned to therapy (any)		Assigned to cash (any)		Both	
			ITT	Std. Err.	ITT	Std. Err.	ITT	Std. Err.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Business investment in past 2 weeks (USD)	2 & 5 wk.	16.812	8.005	[5.468]	57.001	[7.346]***	-17.507	[10000]*
	12 & 13 mo.	26.121	3.123	[12.468]	19.328	[14.144]	-8.919	[18.728]
Estimated value of business assets (USD)	2 & 5 wk.							
	12 & 13 mo.							
Index of income measures (z-score)	2 & 5 wk.	-0.145	0.161	[.087]*	0.318	[.094]***	-0.256	[.123]**
	12 & 13 mo.	-0.016	0.137	[.093]	-0.054	[.090]	-0.056	[.125]
Weekly average profits (USD)	2 & 5 wk.	14.342	1.734	[1.407]	4.518	[1.420]***	-3.287	[2.008]
	12 & 13 mo.	17.591	0.495	[1.567]	1.370	[1.538]	-1.708	[2.196]
Durable assets, z-score	2 & 5 wk.	-0.093	0.128	[.089]	0.151	[.089]*	-0.049	[.127]
	12 & 13 mo.	0.009	0.141	[.089]	-0.113	[.089]	0.058	[.123]
Consumption in the past 2 weeks, USD	2 & 5 wk.	44.146	12.056	[3.576]***	26.570	[3.943]***	-16.582	[5.330]***
	12 & 13 mo.	47.432	-2.290	[3.561]	-2.586	[3.390]	-0.373	[4.321]
Consumption of health and education items in the past 2 weeks, USD	2 & 5 wk.	2.085	3.133	[1.477]**	3.500	[1.651]**	-0.709	[2.369]
	12 & 13 mo.	3.104	-0.897	[.805]	-1.449	[.749]*	1.683	[.963]*
Consumption of food items in the past 3 days, USD	2 & 5 wk.	6.003	1.188	[.383]***	0.963	[.403]**	-1.692	[.555]***
	12 & 13 mo.	6.370	-0.032	[.493]	-0.145	[.445]	-0.541	[.576]
Consumption of non- food items in the past 2 weeks, USD	2 & 5 wk.	12.877	1.495	[1.572]	8.996	[1.688]***	-1.490	[2.346]
	12 & 13 mo.	16.500	-0.908	[2.130]	-1.483	[2.069]	1.155	[2.585]
Savings (USD)	2 & 5 wk.	45.957	-1.169	[8.252]	16.679	[9.260]*	3.219	[12.649]
	12 & 13 mo.	51.395	11.531	[10.270]	2.199	[9.538]	7.996	[13.791]
Number of days they slept on the street in the past 2 weeks	2 & 5 wk.	2.269	0.122	[.271]	-0.527	[.297]*	-0.492	[.379]
	12 & 13 mo.	1.887	0.085	[.284]	-0.095	[.281]	-0.273	[.377]
Sleeping on the streets now	2 & 5 wk.	0.202	0.006	[.027]	-0.091	[.027]***	-0.010	[.035]
	12 & 13 mo.	0.147	0.018	[.028]	0.012	[.027]	-0.049	[.036]
Hours of work per week in past month	2 & 5 wk.	36.376	0.437	[2.666]	6.717	[2.684]**	-5.292	[3.748]
	12 & 13 mo.	34.163	1.351	[2.362]	0.945	[2.329]	-3.540	[3.226]
Hours of illicit work per week	2 & 5 wk.	2.071	-0.649	[.487]	-0.696	[.463]	0.178	[.604]
	12 & 13 mo.	1.482	-0.239	[.431]	0.117	[.418]	-0.335	[.537]
Hours of formal sector work/week	2 & 5 wk.	1.580	0.388	[.449]	0.203	[.425]	-0.052	[.624]
	12 & 13 mo.	2.124	-0.410	[.466]	-0.484	[.423]	0.320	[.578]
Hours of informal/casual work per week	2 & 5 wk.	1.646	0.100	[.157]	0.584	[.162]***	-0.440	[.229]*
	12 & 13 mo.	1.932	0.123	[.225]	0.070	[.236]	-0.184	[.327]
Home robbed, past month	2 & 5 wk.	0.698	-0.028	[.045]	0.002	[.045]	-0.024	[.062]
	12 & 13 mo.	0.711	-0.038	[.043]	-0.055	[.041]	0.077	[.059]
Belongings stolen, past month	2 & 5 wk.	0.784	0.016	[.037]	-0.006	[.038]	-0.051	[.052]
	12 & 13 mo.	0.635	-0.012	[.045]	-0.048	[.046]	0.099	[.062]

Notes: Robust standard errors in brackets, clustered by individual. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table D.8: Robustness of effects of cash grants

Outcome, z-score		Main Specification			Endline Averages, unclustered			Endline Averages, clustered			Conservative Attrition Scenario		
		Therapy Only	Cash Only	Therapy & Cash	Therapy Only	Cash Only	Therapy & Cash	Therapy Only	Cash Only	Therapy & Cash	Therapy Only	Cash Only	Therapy & Cash
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Self control	2–5w	0.087 [.093]	-0.154 [.099]	0.036 [.091]	0.087 [.101]	-0.154 [.107]	0.036 [.098]	0.087 [.162]	-0.154 [.117]	0.036 [.069]	-0.119 [.093]	-0.360 [.098]***	-0.176 [.090]*
	12–13m	0.153 [.082]*	-0.025 [.086]	0.237 [.086]***	0.153 [.089]*	-0.025 [.093]	0.237 [.093]**	0.153 [.106]	-0.025 [.122]	0.237 [.145]	0.010 [.084]	-0.165 [.086]*	0.119 [.087]
Antisocial behaviors	2–5w	-0.197 [.074]***	-0.089 [.073]	-0.254 [.072]***	-0.197 [.080]**	-0.091 [.080]	-0.271 [.078]***	-0.197 [.151]	-0.091 [.087]	-0.271 [.109]*	0.002 [.076]	0.109 [.077]	-0.049 [.075]
	12–13m	-0.092 [.081]	0.090 [.082]	-0.223 [.075]***	-0.081 [.087]	0.106 [.089]	-0.223 [.081]***	-0.081 [.114]	0.106 [.025]***	-0.223 [.142]	0.031 [.079]	0.191 [.080]**	-0.115 [.074]
Violent/criminal attitudes	2–5w	-0.228 [.093]**	-0.225 [.094]**	-0.200 [.098]**	-0.228 [.101]**	-0.225 [.102]**	-0.200 [.106]*	-0.228 [.100]*	-0.225 [.064]**	-0.200 [.095]	0.002 [.091]	0.048 [.093]	0.020 [.095]
	12–13m	-0.079 [.080]	0.014 [.079]	-0.176 [.078]**	-0.073 [.085]	0.023 [.085]	-0.173 [.084]**	-0.073 [.155]	0.023 [.080]	-0.173 [.152]	0.058 [.079]	0.134 [.077]*	-0.049 [.077]
Income	2–5w	0.161 [.087]*	0.318 [.094]***	0.223 [.087]**	0.161 [.094]*	0.318 [.102]***	0.223 [.094]**	0.161 [.079]	0.318 [.136]*	0.223 [.071]**	-0.050 [.087]	0.049 [.093]	-0.005 [.085]
	12–13m	0.137 [.093]	-0.054 [.090]	0.027 [.089]	0.137 [.100]	-0.054 [.097]	0.027 [.096]	0.137 [.109]	-0.054 [.137]	0.027 [.127]	0.008 [.091]	-0.161 [.089]*	-0.082 [.088]