Supplementary Materials

Yield Model R-code and Output:

```
Family: gaussian
Link function: identity

Formula:
HeavyCrop ~ s(Cocoa.density, k = 5, bs = "tp") + s(Canopy.gap.dry,
   k = 5, bs = "tp") + s(PropCPB, k = 5, bs = "tp") + s(soil.moist,
   k = 5, bs = "tp") + s(Biomass, k = 5, bs = "tp") + s(No.applications.yr,
   k = 5, bs = "tp") + s(distance.cont, k = 5, bs = "tp")

Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.7413     0.0366   20.25   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
                    edf Ref.df     F  p-value
s(Cocoa.density)     0.9304      4 3.341 0.000518 ***
s(Canopy.gap.dry)    0.6312      4 0.428 0.107842
s(PropCPB)           0.8148      4 1.100 0.024433 *
s(soil.moist)        0.6147      4 0.251 0.198157
s(Biomass)           0.9463      4 4.406  9.2e-05 ***
s(No.applications.yr) 0.9419     4 4.050 0.000231 ***
s(distance.cont)     0.8333      4 1.250 0.017192 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =   0.52   Deviance explained = 59.9%
-REML = 5.6525  Scale est. = 0.04823   n = 36
```

#Cocoa income Relationships to Poverty Measures
Relationship between cocoa income as continuous and education outcome (likelihood child misses school)

```
x<-glm(Education1~Cocoa.Income,family=binomial,data=dF.pov)
summary(x)

Call:
glm(formula = Education1 ~ Cocoa.Income, family = binomial, data = dF.pov)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.45909  -1.20136  -0.03936   1.02264   1.67223

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)   6.415e-01  2.824e-01   2.272  0.02312 *
```

```
Cocoa.Income -2.787e-04  9.663e-05  -2.884  0.00392 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

   Null deviance: 148.32  on 106  degrees of freedom
Residual deviance: 134.35  on 105  degrees of freedom
AIC: 138.35

Number of Fisher Scoring iterations: 5
```

With cocoa income quartile

```
                    Df Sum Sq Mean Sq F value Pr(>F)
factor(Cocoa.income.quart)  3  2.614  0.8713   3.732 0.0137 *
Residuals                  99 23.114  0.2335
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
4 observations deleted due to missingness
```

Tukey multiple comparisons of means
95% family-wise confidence level

```
Fit: aov(formula = glm(Education ~ factor(Cocoa.income.quart), family = binomial, data = dF.pov))

$`factor(Cocoa.income.quart)`
        diff        lwr       upr     p adj
2-1 0.02615385 -0.32753772 0.3798454 0.9974262
3-1 0.24000000 -0.11714230 0.5971423 0.3007355
4-1 0.38370370  0.03323755 0.7341699 0.0260740
3-2 0.21384615 -0.13984542 0.5675377 0.3946945
4-2 0.35754986  0.01060083 0.7044989 0.0407824
4-3 0.14370370 -0.20676245 0.4941699 0.7076261
```

Relationship between cocoa income as a continuous variable and probability of owning a TV as proxy for asset acquisition

```
x<-glm(TV~Cocoa.Income,family=binomial,data=dF.pov)
summary(x)

Call:
glm(formula = TV ~ Cocoa.Income, family = binomial, data = dF.pov)

Deviance Residuals:
   Min      1Q  Median      3Q     Max
-2.2144  -0.8861  -0.7955   1.1837   1.6318

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)  -1.025e+00  2.815e-01  -3.640 0.000273 ***
Cocoa.Income  2.016e-04  7.672e-05   2.627 0.008604 **
```

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 142.44 on 106 degrees of freedom
Residual deviance: 131.39 on 105 degrees of freedom
AIC: 135.39

Number of Fisher Scoring iterations: 4

With cocoa income quartile

                        Df Sum Sq Mean Sq F value Pr(>F)
factor(Cocoa.income.quart)   3 2.664 0.8879   4.042 0.00923 **
Residuals   103 22.626 0.2197

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Tukey multiple comparisons of means
95% family-wise confidence level

Fit: aov(formula = glm(TV ~ factor(Cocoa.income.quart), family = binomial, data = dF.pov))

$`factor(Cocoa.income.quart)`
        diff   lwr   upr   p adj
2-1 0.14102564 -0.19528492 0.4773362 0.6933376
3-1 0.16769231 -0.17515405 0.5105387 0.5793896
4-1 0.42838196 0.09780941 0.7589545 0.0055010
3-2 0.02666667 -0.31305317 0.3663865 0.9969337
4-2 0.28735632 -0.03997251 0.6146852 0.1064415
4-3 0.26068966 -0.07335075 0.5947301 0.1809176

Relationship between cocoa income quartile and reported food security

x<-glm(Food.amount~Cocoa.income.quart,family=binomial,data=dF.pov)
summary(x)

Call:
glm(formula = Satisfaction.life.overall ~ Cocoa.income.quart,
    family = poisson, data = dF.pov)

Deviance Residuals:
    Min   1Q   Median   3Q   Max
-1.8974 -0.4033 -0.1186   0.5305   1.3037

Coefficients:
                  Estimate Std. Error z value Pr(>|z|)
(Intercept)   0.2384   0.1741   1.370   0.1708
Cocoa.income.quart2   0.2724   0.2292   1.189   0.2346
Cocoa.income.quart3   0.3494   0.2292   1.524   0.1274

Cocoa.income.quart4 0.5374 0.2149 2.501 0.0124 *

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 72.416 on 106 degrees of freedom
Residual deviance: 65.735 on 103 degrees of freedom
AIC: 313.01

Number of Fisher Scoring iterations: 5

Tukey multiple comparisons of means
95% family-wise confidence level

Fit: aov(formula = glm(Satisfaction.life.overall ~ factor(Cocoa.income.quart), family = poisson, data = dF.pov))

$`factor(Cocoa.income.quart)`

	diff	lwr	upr	p adj
2-1	0.3974359	-0.2319904	1.0268622	0.3561118
3-1	0.5307692	-0.1108892	1.1724277	0.1414344
4-1	0.9031830	0.2844958	1.5218703	0.0013241
3-2	0.1333333	-0.5024736	0.7691403	0.9469890
4-2	0.5057471	-0.1068693	1.1183635	0.1426911
4-3	0.3724138	-0.2527638	0.9975913	0.4085064

Figures

Figure S1 Ecological Production Function Diagnostic Plots To confirm that (top row) residuals are normally distributed, (bottom left) residuals versus fitted values are randomly distributed and (bottom right) fitted values are reasonable approximations of response values .

Figure S2 Relationship Between Yield and Input/Labour Costs Linear regression of input (left) and labour (right) costs with monitored cocoa yields, to generate predictive relationships of management costs for enhanced yields. Per hectare cost estimates were derived from household survey responses.

Figure S3 Farm Characteristics Over Distance from Forest To assess whether farms showed significant differences in (a) age of cocoa, (b) available soil phosphorus, (c) soil carbon to nitrogen ratios or (d) soil potassium with distance from forest. If these were significant, it would suggest yield benefits of proximity to forest would be driven by younger more vigorous farms being located at the forest frontier.

Figure S4 Potential Yield Increase per Farm Modelled per hectare yield increases in decreasing order by farm. Farm yields per factor were calculated using our cocoa ecological production function, field-measured minimum or maximum values for the three chosen factors and farm-specific conditions. Hollow bars represent potential yields and grey shaded bars represent current yields.

Figure S5 Shifts in Probability of Poverty Outcomes with Increased Incomes Modelled probabilities of per household poverty outcome plotted over the log of potential income for each management intervention. These distributions illustrates shifts in the probability for (a) a child misses school and (b) a household is able acquire assets like a TV from original incomes (grey points) to estimated higher incomes (black points). Dashed lines indicate mean probabilities for the original study conditions (grey dotted line) and for probabilities under each management factor (black dotted lines).

Figure S6 Shifts in Probability of Poverty Outcomes with Changing Income Quartiles Modelled changes in probability of a household being food secure from original conditions (grey points) and under higher yields for each management intervention (black points) ordered by potential income. Dashed lines indicate mean probabilities for the original study conditions (grey dotted line) and for probabilities under each management factor (black dotted lines).

Tables

Table S1 Chosen indicators for ecological processes of interest for the ecological production function.

Ecological Process	Indicator	Description
Micro-climate	Water Stress	Measure, in mm, of difference between vegetation water demand, precipitation and infiltration averaged over growing and harvesting season.
	Maximum Temperature	Maximum temperature, in Celsius, over growing and harvesting season. Measured at 30 min intervals within plots.
Pollination	Cherelle set	Percent of monitored flowers and flower buds that are pollinated and form cherelles.
Disease	Mistletoe	Percent of monitored disease trees with mistletoe present in canopy.
	Capsid Attack	Percent of monitored pods with visible capsid attack.
	Black Pod	Percent of monitored pods with visible black pod infection.
Soil	C:N Ratio	Ratio of carbon and nitrogen content measured in top 0-30 cm of soils.
	Phosphorus	Available phosphorus, in ppm, measured in top 0-30 cm of soils
	pH	Soil pH measured in top 0-30 cm.
	Soil Moisture	Volumetric water content measured monthly next to monitored to disease trees, averaged over the plot and the growing and harvesting season.
Off-Farm ES	Distance from Forest	Continuous measure, in m, of established plot from forest edge, derived from Landsat imagery.
Management	Cocoa Density	Number of cocoa trees planted, per hectare.
	Canopy Gap	Percent openness of shade tree canopy.
	Distance to Rotting Biomass	Average distance, in m, of closest area of rotting biomass to monitored fruitset trees.
	Density of Shade Trees	Number of non-cocoa trees, per hectare, left in the farm that are at least 12 m tall.
	Age of Cocoa	Farmer reported age of cocoa.
	Fertilizer Application	Average application rate, per year, derived from survey asking number of total times over previous 5 years.

Table S2 ANOVA analysis of household poverty measures relative to cocoa income and income quartiles.

Dimension	Indicator	Description	Cocoa Income	Cocoa Income Quartiles
Health	Under 5 mortality	Binary variable. 1 means the household has not ever experienced.	NS	NS
	Perceived adequateness of access to health care	4 point Likert scale in response to question 'Does the household have adequate access to healthcare?'. Higher numbers correspond to agreement.	NS	NS
Education	Household head literacy	Binary variable. 1 means the household head is literate.	NS	NS
	Child missed school in the last year	Binary variable. 1 means the child has missed school because household could not afford costs.	**	Q4-Q1 ** Q4-Q2 *
Basic needs	Electricity	Binary variable. 1 if household has access to electricity.	NS	NS
	Access to improved sanitation	Binary variable. 1 if household has access to improved sanitation	NS	NS
	Access to clean water	Binary variable. 1 if household has access to clean drinking water within 30 minute walk.	NS	NS
Assets	TV	Binary variable. 1 if the household owns a TV.	**	Q4-Q1 **
Satisfaction	Satisfaction with life overall	4 point Likert scale. Higher numbers correspond to high satisfaction.	NS	Q4 *
Food security	Adequate amount of food in the last year	Binary variable. 1 means the household had 0 months without enough food.	NS	Q4-Q1 **
	Adequate variety of food in the last year	Binary variable. 1 means the household had 0 months without an adequate variety of food.	.	NS
Empowerment	Could easily access more land	4 point agree-disagree Likert scale in response to statement 'I could easily get access to more land if I wanted to'. Higher numbers correspond to agreement.	NS	NS
Social connectedness	Access to extension in the last 2 years	Binary variable. 1 means the household had received (state or private) agricultural extension/training in the last 2 years.	NS	NS

NS not significant, . <.1, * <.05, ** <.01,***<.001
Q1=poorest, *Q4*=richest.