The Farmer and the Fates: Locus of Control and Investment in Rainfed Agriculture

JG Malacarne
University of Maine

Supplemental Materials

A Data Collection Timeline

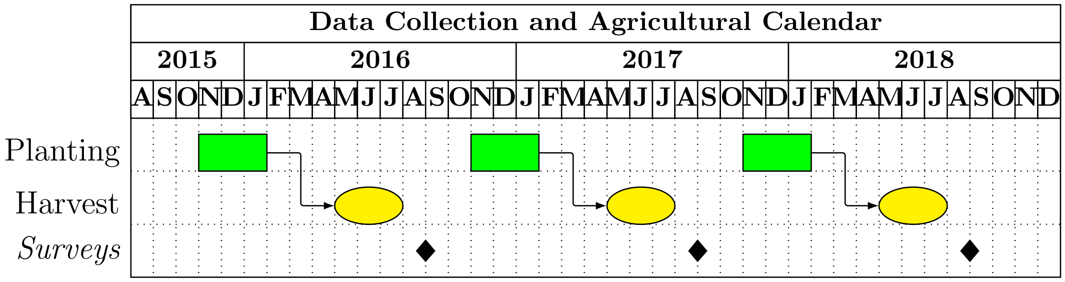

Figure A.1: Household characteristics and beliefs from each survey wave are used to describe investment choices in the following planting season, which occurred a few months later.

B Analytical Solution

Consider a production technology that generates output from two inputs (x,e). Let x be an input under the decision-maker's control and let e be an input that is not. Assume that the decision-maker can describe her beliefs about the possible realizations of e using the probability density function $\phi_i(e)$. Given an input cost (c), output price (p) and wealth endowment (A_i), the expected utility maximization problem can be written as:

$$\max_x E[U_i] = E\left[U_i\left(pf_i(x,e) + A_i - cx\right)\right] \quad s.t. \quad x \geq 0,\ cx \leq A_i \tag{8}$$

Assume a concave utility function and positive but diminishing marginal returns to each input in $f(x,e)$. Writing the Lagrangian associated with the maximization program:

$$\mathcal{L} = E\left[U_i(pf_i(x,e) + A_i - cx)\right] - \lambda(cx - A_i) + \lambda_x x$$

The first order conditions of which are:

$$\begin{aligned} &(1) \quad E\left[U_i'(\cdot)(pf_{ix} - c)\right] - \lambda c + \lambda_x = 0 \\ &(2) \quad \lambda(cx - A) = 0 \\ &(3) \quad \lambda_x x = 0 \end{aligned}$$

And the second order condition:

$$\frac{\partial^2 \mathcal{L}}{\partial x_{it}^2} = E\left[U_i'(\cdot)pf_{ixx} + U_i''(\cdot)pf_{ix}\right] < 0$$

Rearranging the second order condition yields:

$$\underbrace{E[\underbrace{U_i'(\cdot)}_{+}\underbrace{f_{ixx}}_{-}]}_{-} < \underbrace{-E[\underbrace{U_i''(\cdot)}_{-}\underbrace{f_{ix}}_{+}]}_{+} \tag{9}$$

Given the assumptions on $U(\cdot)$ and $f(x,e)$, the left-hand side of Equation 9 is negative and the right-hand side is positive. The second order condition is thus satisfied everywhere.

As discussed in Section 3, let locus of control for individual i be denoted by LOC_i. Keeping with convention in psychology, larger values of LOC_i are associated with a more external locus of control. Let Z_i be a vector of other individual characteristics that might affect beliefs about the production function – such as location or ability. The production function can then be written as:

$$f_i(x, e; Z_i, LOC_i) \tag{10}$$

In this notation variable inputs appear before the semi-colon and parameters influencing the shape of the production function after the semi-colon. Assume that marginal products are positive for both inputs and exhibit diminishing marginal returns.

$$\frac{\partial f_i(x,e;Z_i,LOC_i)}{\partial x} > 0 \qquad \frac{\partial^2 f_i(x,e;Z_i,LOC_i)}{\partial x^2} < 0$$
$$\frac{\partial f_i(x,e;Z_i,LOC_i)}{\partial e} > 0 \qquad \frac{\partial^2 f_i(x,e;Z_i,LOC_i)}{\partial e^2} < 0$$

Also assume that the two inputs are complements in production.

$$\frac{\partial^2 f_i(x,e;Z_i,LOC_i)}{\partial x \partial e} > 0 \qquad \frac{\partial^2 f_i(x,e;Z_i,LOC_i)}{\partial e \partial x} > 0 \tag{11}$$

Section 3 posits that a more external locus of control will reduce the decision-maker's perception of the return to her own investment and increase her perception of the return to the input she does not control. This is captured in the model by the assumed signs on the partial derivatives below, once again noting that LOC_i is a parameter and not a variable input in the production function.

$$\frac{\partial^2 f_i(x,e;Z_i,LOC_i)}{\partial LOC_i\, \partial x} < 0 \qquad \forall\, x \tag{12}$$
$$\frac{\partial^2 f_i(x,e;Z_i,LOC_i)}{\partial LOC_i\, \partial e} > 0 \qquad \forall\, e \tag{13}$$

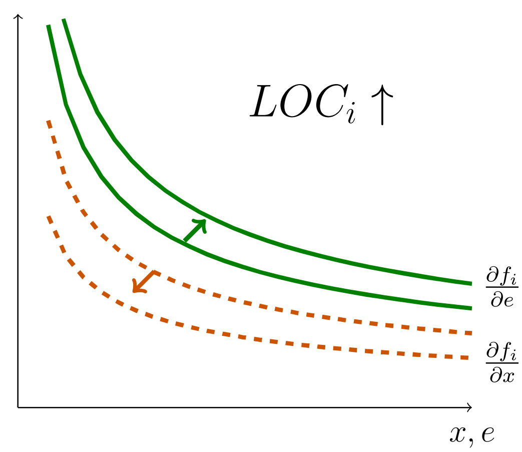

Figure A.2: A more external locus of control is hypothesized to increase the perceived marginal return to e and decrease the perceived marginal return to x.

To see how locus of control affects investment decisions in the interior of the optimization program, consider the first order condition (1).

$$E\left[U_i'(\cdot)(pf_{ix})\right] - E\left[U_i'(\cdot)(c)\right] = 0$$

Rearranging terms and expanding the expectation

$$p \int_e \left(U_i'(pf_i(x,e) + A_i - cx) \frac{\partial f_i(x,e)}{\partial x} \right) \phi_i(e) de = c \int_e \left(U_i'(pf_i(x,e) + A_i - cx) \right) \phi_i(e) de \tag{14}$$

Note that the marginal product of the externally controlled input does not appear in Equation 14. The influence of locus of control on investment will thus be driven by the changes it induces on the perceived marginal product of x. Rearranging equation 14:

$$\frac{p}{c} = \frac{\int_e U_i'(pf_i(x,e) + A_i - cx)\phi_i(e)de}{\int_e U_i'(pf_i(x,e) + A_i - cx)\frac{\partial f_i(x,e)}{\partial x}\phi_i(e)de} \tag{15}$$

Given diminishing marginal returns in x, the right-hand side of Equation 15 approaches p/c from below. A more external locus of control implies a reduction in $\frac{\partial f(x,e)}{\partial x}$ at every value of x and thus an increase in the right-hand side of Equation 15. Equation 15 will thus be satisfied at a lower level of investment.

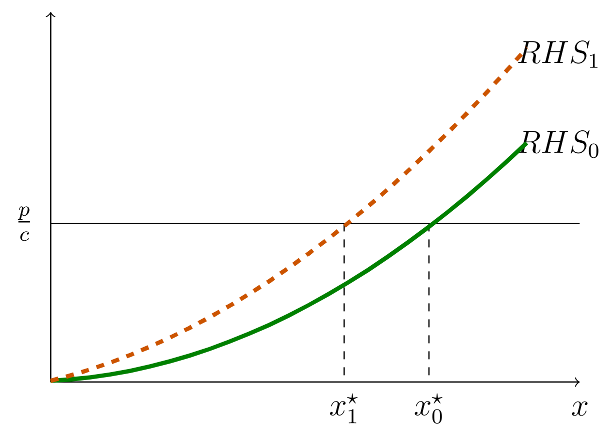

Figure A.3: A more external locus of control decreases $\partial f_i/\partial x$ and shifts the RHS of Equation 15 upward at any value of x.

B.1 Zero Investment Corner

First consider the possibility that $x^\star = 0$. Let $\lambda_x > 0$. Then by FOC (3) $x = 0$. At $x = 0$, FOC (2) becomes $\lambda(-A) = 0$ and thus $\lambda = 0$

FOC (1) then becomes:

$$E\left[U_i'(\cdot)(pf_{ix} - c)\right] < 0$$
$$E\left[U_i'(\cdot)(pf_{ix})\right] < E\left[U_i'(\cdot)c\right] \tag{16}$$

Equation 16 says that, for zero investment to be optimal, the utility gain from the value marginal product of x must be less than the utility cost of investment as x approaches zero. Reorganizing slightly:

$$\frac{E\left[U_i'(\cdot)(f_{ix})\right]}{E\left[U_i'(\cdot)\right]} < \frac{c}{p}$$

This condition does not hold. The right-hand side is a constant. By assumption, the marginal product of $f()$ with respect to x gets arbitrarily large as x approaches zero. Thus while failure to adopt a technology will be of significant in sections discussing adoption, it will not come from the solution to the expected utility maximization problem occurring at the zero investment corner. It is worth nothing, however, that if the input is at all bulky then the optimal investment could be smaller than the available unit, resulting in zero investment. Agents solving the expected utility maximization problem and comparing across different technologies options could also fail to adopt.

B.2 Liquidity Constrained Corner

One of the primary populations of interest in studying rainfed agriculture is small-holding households producing primarily for subsistence purposes. As such, it is likely that many households make decisions in the proximity of liquidity constraints. While broader issues of access will also affect technology adoption, liquidity constraints will certainly play a significant role in investment decisions.

Let $\lambda > 0$, then by FOC (2) $x^{\star} = \frac{A}{c}$. With $x > 0, \lambda_x = 0$ by FOC (3). At $x^{\star} = \frac{A}{c}$, FOC (1) then becomes:

$$E\left[U_i'(\cdot)(pf_{ix} - c)\right]|_{x=\frac{A}{c}} > 0$$

$$E\left[U_i'(\cdot)(pf_{ix})\right]|_{x=\frac{A}{c}} > E\left[U_i'(\cdot)c\right]|_{x=\frac{A}{c}} \tag{17}$$

This condition is the opposite of the zero investment condition. The liquidity constraint binds if the utility gain from the value marginal product of x is still greater than the utility cost of investment when all available resources are being used for investment. The level of liquidity, $\underline{A}$, an agent needs to avoid being constrained in her choice of $x_{it}^{\star}$ is thus given by the cost of the input multiplied by the profit maximizing input level. That is to say, the level of A necessary to avoid being liquidity constrained moves opposite the optimal investment level. For a given wealth distribution, increasing the desired investment level also increases the chance that a household risks all of its resources on the production process. This is intuitive, but worth noting.

C Using Factor Analysis to Generate a General Locus of Control Measure

To construct a measure of general locus of control, I begin with the twenty-one Likert-style items listed in Table A.1. Seven items pertain to each of three hypothesized control dimensions: internal control, chance factors, and powerful other agents. In response to each item, a respondent reports whether she strongly disagrees, weakly disagrees, weakly agrees, or strongly agrees. Her response is recorded on a scale from 1 (strongly disagree) to 4 (strongly agree).

I begin by correcting each item for acquiescence bias. Following Rammstedt, Kemper and Borg (2013) and Laajaj and Macours (2021), I calculate an acquiescence score for each individual reflecting the difference in her propensity to agree and disagree. Individuals who agree much more often than they disagree have a higher acquiescence score. The acquiescence score is then subtracted from each of the 21 items. The corrected items reflect an individual's response to each item relative to her average response across all items.

After correcting for acquiescence bias, I begin an exploratory factor analysis procedure to determine the structure of the latent variable, or variables, underlying responses to the general locus of control items. I use Stata's *factor* routine with oblique quartimin rotation for all factor analysis. The oblique rotation acknowledges potential correlation across factors. This would allow, as we might expect, factors representing internal and external dimensions of control to be negatively correlated.

Figure A.4 contains a screeplot of the eigenvalues resulting from the first round of exploratory factor analysis. Two criteria are typically used to decide how many factors to retain. The first, called the Kaiser criterion, suggests retaining factors whose eigenvalues are greater than one. The second criterion, called the "elbow criterion", suggests identifying the point where the scree plot flattens dramatically, indicating diminishing returns to additional factors in explaining variation in the data. Both criteria suggest retaining a large number of factors and Stata returns a warning of a Heywood case, suggesting collinearity among some of the items.

In order to address both the Heywood case and the lack of a consistent factor structure, I conduct additional rounds of factor analysis on the set of items associated with each Levenson dimension. In each round, I retain items that load positively on the primary factor and discard those that do not. Visual inspection of the correlation matrix confirms that the discarded items were not correlated with other items in their own scale or with items in the other two scales.[15] In total, fourteen items were retained.

I once again run the factor analysis routine and consider the resulting screeplot in order to asses the underlying factor structure. The Heywood case has been resolved and a Kaiser-

[15] This type of selection is consistent with the design of psychometric instruments and with locus of control. Rotter (1975) writes, "A theorist may choose to use a construct of any breadth that he wishes, as long as it meets the criterion of functionality. That is, the referents that are included within the construct have a greater than chance correlation. Not every referent must correlate greater than chance with every other, but any referent must on the average correlate better than chance with all of the others. This is the same criteria that should be used in developing a measure of the same construct; namely, that each item should correlate significantly with the sum of the other items, with that item removed."

Meyer-Olkin test (0.73) indicates that there is a sufficient common factor structure to warrant data reduction via factor analysis. Both selection criteria now agree that a single factor should be retained (Figure A.5). Plotting the resulting factor loadings in Figure A.6 also indicates a clear pattern across internal items (negative loadings) and external items (positive loadings). I interpret this factor as “general locus of control”. In keeping with psychology convention, larger values are associated with more external locus of control.

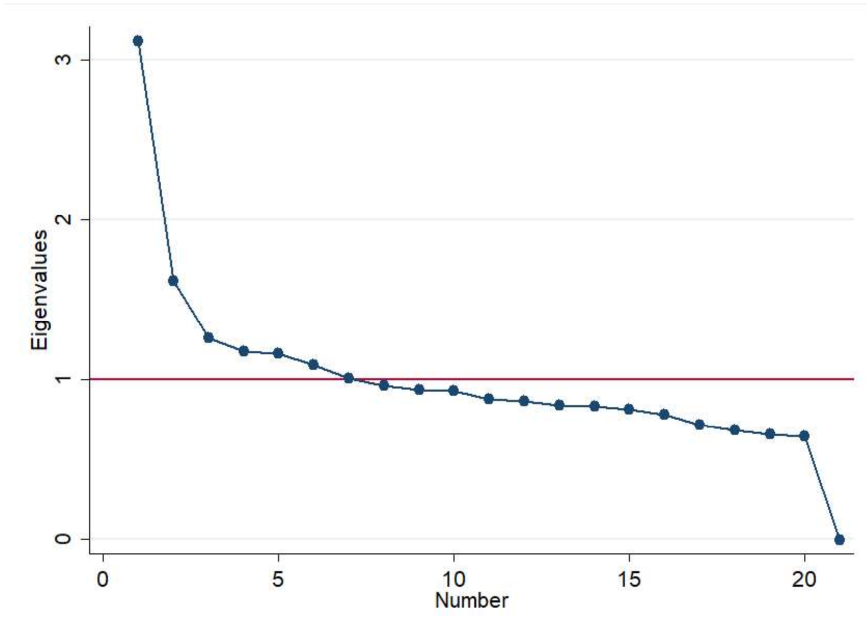

Figure A.4: Screeplot - All General Locus of Control Items

Table A.1: General Locus of Control Scale Adapted from Levenson (1981) for use in Tanzania and Mozambique

Retained	Internal (I)
X	1. Whether or not I get to be a leader depends mostly on my ability.
X	2. When I make plans, I am almost certain to make them work.
X	3. How many friends I have depends on how nice a person I am.
X	4. I can pretty much determine what will happen in my life.
X	5. I am usually able to protect my personal interestes.
X	6. When I get what I want, it's usually because I worked hard for it.
X	7. My life is determined by my own actions.
Retained	**Chance (C)**
X	1. To a great extent my life is controlled by accidental happenings.
	2. Often there is no chance of protecting my personal interests ... from bad luck happenings.
X	3. When I get what I want, it's usually because I'm lucky.
	4. I have often found that what is going to happen will happen.
	5. It is not always wise for me to plan too far ahead because many... things turn out to be a matter of good or bad fortune.
X	6. Whether or not I get to be a leader depends on whether ... I'm lucky enough to be in the right place at the right time.
	7. It is chiefly a matter of fate ... whether or not I have a few friends or many friends.
Retained	**Powerful Others (P)**
X	1. I feel like what happens in my life... is mostly determined by powerful people.
	2. Although I might have good ability, I will not be given ... leadership responsibility without appealing to those in positions of power.
X	3. My life is chiefly controlled by powerful others.
	4. People like myself have very little chance of protecting our personal ... interests when they conflict with those of strong pressure groups.
X	5. Getting what I want requires pleasing those people above me.
	6. If important people were to decide they didn't like me... I probably wouldn't make many friends.
X	7. In order to have my plans work, I make sure that they fit in ... with the desires of people who have power over me.

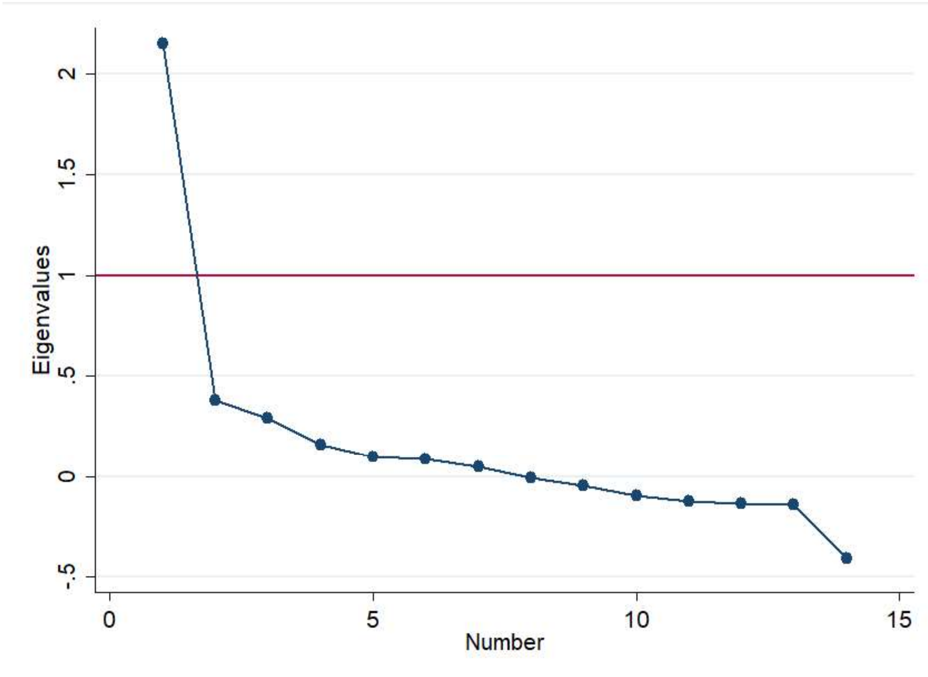

Figure A.5: Screeplot - Retained General Locus of Control Items

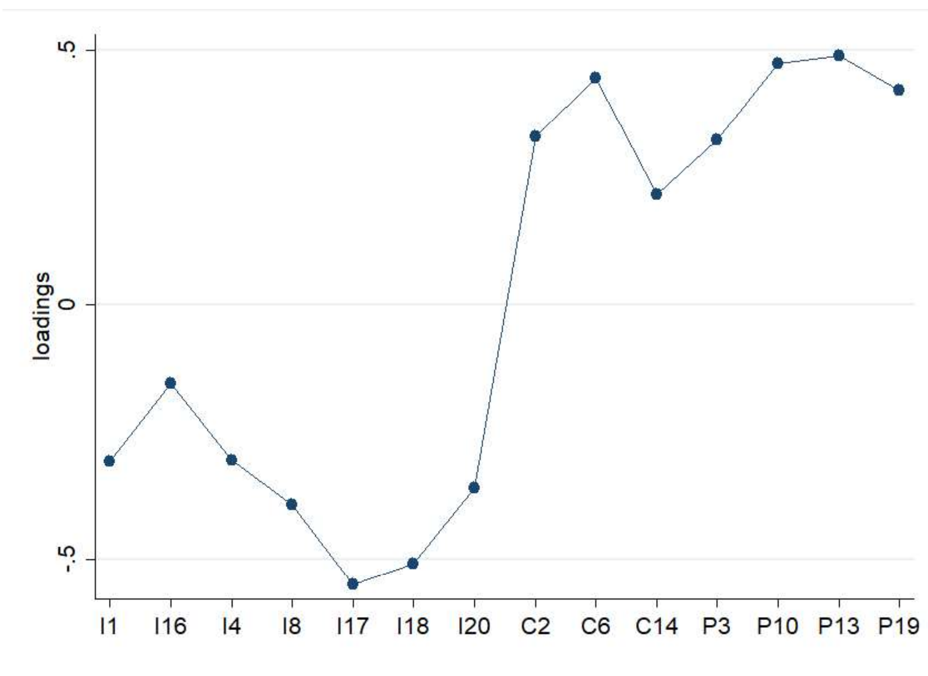

Figure A.6: Factor Loadings - Retained General Locus of Control Items

D Maize-specific Locus of Control

In Section 4.2, maize-specific locus of control is presented as the ratio of a standardized measure of variability attributed to external influences ((V_E) to a standardized measure of variability attributed to internal influences (V_I).

$$LOC_{Maize} = \frac{V_E}{V_I}$$

More specifically, let the three input bundles presented to the respondent be given by $j \in \{1, 2, 3\}$ and the three states of weather be given by $k \in \{l, n, g\}$ (low, normal, good).The harvest expected from using input bundle j in weather state k is then y_{jk}. Average values across weather states, holding input bundles fixed, will replace the k subscript with a bullet point (as in: $\bar{y}_{j\bullet}$). Likewise, averages across input bundles, holding weather state fixed, will replace the j subscript with a bullet point (as in: $\bar{y}_{\bullet k}$).

Maize-specific locus of control is then defined as:

$$LOC_{Maize} = \frac{V_E}{V_I} = \frac{\sum_j \left(\frac{S_{j\bullet}}{\bar{y}_{j\bullet}}\right)}{\sum_k \left(\frac{S_{\bullet k}}{\bar{y}_{\bullet k}}\right)} \tag{18}$$

Where

$$S_{j\bullet} = \sqrt{\frac{\sum_k (y_{jk} - \bar{y}_{j\bullet})^2}{2}} \quad , \quad S_{\bullet k} = \sqrt{\frac{\sum_j (y_{jk} - \bar{y}_{\bullet k})^2}{2}}$$

$$\bar{y}_{j\bullet} = \frac{1}{3}(y_{jl} + y_{jn} + y_{jg}) \quad , \quad \bar{y}_{\bullet k} = \frac{1}{3}(y_{1k} + y_{2k} + y_{3k})$$

E Expected Return to Improved Maize Varieties

As in Section 4.3, let $f_{i,L}$, $f_{i,N}$, and $f_{i,G}$ represent the increase in yield from planting improved maize seed rather than local maize seed as illustrated in Figure A.7.

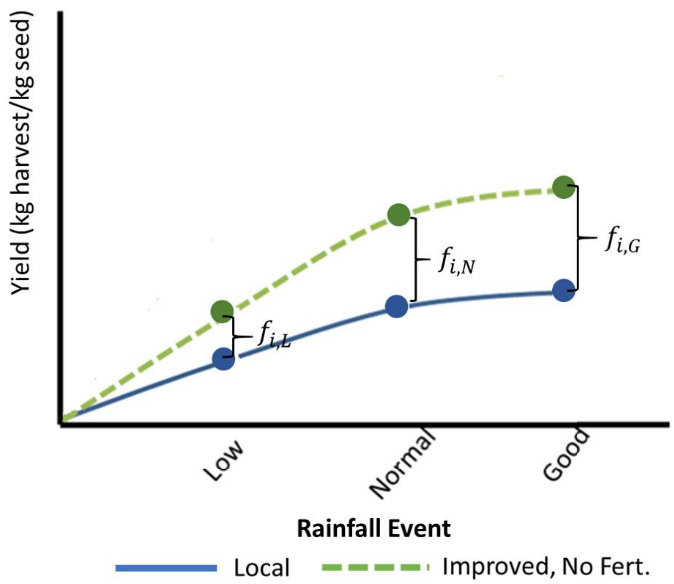

Figure A.7: Expected Yield Gain From Improved Maize Seed

The expected return to using improved maize varieties over local maize varieties can then be calculated as:

$$E[f_{it}] = p_{i,L}f_{i,L} + p_{i,N}f_{i,N} + p_{i,G}f_{i,G}$$

Two moving pieces in the equation above — subjective beliefs about the returns to improved maize seed in a given state of nature and subjective beliefs regarding the probability of being in that state of nature. Recent experience with drought conditions are likely to affect their salience to the decision-maker, and thus subjective probabilities. In the conceptual model presented in Section 3, this would represent a separate influence from locus of control. As such, I fix each probability at 1/3 for all the analysis in this paper.

The table below shows rainfall expectations of households in the sample. On average, they expect one more normal year than good year.

The results in Section 6 reporting the relationship of locus of control and expected returns are robust to using subjective weights rather than equal weights. The number of years of good rainfall and the number of years of normal rainfall are also included as controls in all specifications in Table 4 that note the use of "Additional Controls".

Table A.2: Subjective Rainfall Expectations

Expected Years (Out of 10)	Mozambique	Tanzania	Total
Low Rainfall Years	2.4	2.9	2.7
Normal Rainfall Years	4.4	4.0	4.2
Good Rainfall Years	3.2	3.1	3.1

F Maize-specific Locus of Control is not mechanically linked to expected return to investment (Proof)

In this section, I show that, despite being constructed with the same data, the expected yield gain and maize-specific locus of control measure are conceptually and empirically distinct. In summary, I show that locus of control is invariant to scaling all outcomes by a positive constant, while expected return is not. Adding a positive constant to all outcomes, on the other hand, does not change the expected return but does change the maize-specific locus of control measure.

Consider the following example with two possible choices by the decision-maker $I \in \{1, 2\}$ and two states of nature, $E \in \{1, 2\}$.

		E: 1	E: 2
I	1	y_{11}	y_{12}
	2	y_{21}	y_{22}

Assume a positive return to the decision-maker's (DM) choice of $I = 2$ in both states of nature. Also let output for both choices by the DM increase if the state of nature is $E = 2$. No restriction is placed on the relationship between y_{21} and y_{12}. Thus the initial conditions are:

$$\begin{array}{ll} y_{11} < y_{21} & y_{11} < y_{12} \\ y_{12} < y_{22} & y_{21} < y_{22} \end{array}$$

Assigning any probabilities (P_1,P_2) to the states of nature, the expected return to adoption is given by

$$E[R_0] = P_1(y_{21} - y_{11}) + P_2(y_{22} - y_{12}) \tag{19}$$

As in Section 4.2, define activity specific locus of control for $j = I \in \{1, 2\}$ and $k = E \in \{1, 2\}$ as:

$$LOC_0 = \frac{\sum_j \left(\frac{S_{j\bullet}}{\bar{y}_{j\bullet}}\right)}{\sum_k \left(\frac{S_{\bullet k}}{\bar{y}_{\bullet k}}\right)}$$

Where

$$S_{j\bullet} = \sqrt{\sum_k (y_{jk} - \bar{y}_{j\bullet})^2}$$

$$S_{\bullet k} = \sqrt{\sum_j (y_{jk} - \bar{y}_{\bullet k})^2}$$

$$\bar{y}_{j\bullet} = \frac{1}{2}(y_{j1} + y_{j2})$$

$$\bar{y}_{\bullet k} = \frac{1}{2}(y_{1k} + y_{2k})$$

Given the simple two-by-two space. I write LOC_0 as a single expression made up of four elements.

$$LOC_0 = \frac{\frac{\sqrt{(y_{11}-\bar{y}_{1\bullet})^2+(y_{12}-\bar{y}_{1\bullet})^2}}{\bar{y}_{1\bullet}} + \frac{\sqrt{(y_{21}-\bar{y}_{2\bullet})^2+(y_{22}-\bar{y}_{2\bullet})^2}}{\bar{y}_{2\bullet}}}{\frac{\sqrt{(y_{11}-\bar{y}_{\bullet 1})^2+(y_{21}-\bar{y}_{\bullet 1})^2}}{\bar{y}_{\bullet 1}} + \frac{\sqrt{(y_{12}-\bar{y}_{\bullet 2})^2+(y_{22}-\bar{y}_{\bullet 2})^2}}{\bar{y}_{\bullet 2}}} \tag{20}$$

Where:

$$\begin{array}{ll} \bar{y}_{1\bullet} = \frac{1}{2}(y_{11} + y_{12}) & \bar{y}_{\bullet 1} = \frac{1}{2}(y_{11} + y_{21}) \\ \bar{y}_{1\bullet} = \frac{1}{2}(y_{21} + y_{22}) & \bar{y}_{\bullet 2} = \frac{1}{2}(y_{12} + y_{22}) \end{array}$$

From left to right and top to bottom, call the four elements in Equation 20 $N_{1\bullet}$, $N_{2\bullet}$, $N_{\bullet 1}$, and $N_{\bullet 2}$

F.1 Scaling outcomes by a constant

Consider

		E: 1	E: 2
I	1	y_{11}^{λ}	y_{12}^{λ}
	2	y_{21}^{λ}	y_{22}^{λ}

where $y_{jk}^{\lambda} = \lambda \cdot y_{jk}$

F.1.1 Scaling all outcomes by a constant changes the expected return to adoption.

$$\begin{aligned} E[R^{\lambda}] &= P_1(y_{21}^{\lambda} - y_{11}^{\lambda}) + P_2(y_{22}{}^{\lambda} - y_{12}{}^{\lambda}) \\ &= P_1(\lambda y_{21} - \lambda y_{11}) + P_2(\lambda y_{22} - \lambda y_{11}) \\ &= \lambda P_1(y_{21} - y_{11}) + \lambda P_2(y_{22} - y_{11}) \\ &= \lambda E[R_0] \neq E[R_0] \end{aligned}$$

F.1.2 Scaling all outcomes by a constant does not change activity specific locus of control

There are four elements in LOC_{λ}. From left to right and top to bottom, call these four elements $N_{1\bullet}^{\lambda}$, $N_{2\bullet}^{\lambda}$, $N_{\bullet 1}^{\lambda}$, and $N_{\bullet 2}^{\lambda}$. The numerator and denominator of each element is of the same form. Consider the numerator of the first element.

$$\sqrt{(y_{11}^{\lambda} - \bar{y}_{1\bullet}^{\lambda})^2 + (y_{12}^{\lambda} - \bar{y}_{1\bullet}\lambda)^2} \tag{21}$$

Note that:

$$\begin{aligned} \bar{y}_{1\bullet}^{\lambda} &= \frac{1}{2}(y_{11}^{\lambda} + y_{12}^{\lambda}) \\ &= \frac{1}{2}(\lambda y_{11} + \lambda y_{12}) \\ &= \frac{1}{2}\lambda(y_{11} + y_{12}) = \lambda \bar{y}_{1\bullet} \end{aligned}$$

Then 21 becomes:

$$\begin{aligned} &\sqrt{(\lambda y_{11} - \lambda \bar{y}_{1\bullet})^2 + (\lambda y_{12} - \lambda \bar{y}_{1\bullet})^2} \\ &= \sqrt{\lambda^2(y_{11} - \bar{y}_{1\bullet})^2 + \lambda^2(y_{12} - \bar{y}_{1\bullet})^2} \\ &= \lambda\sqrt{(y_{11} - \bar{y}_{1\bullet})^2 + (y_{12} - \bar{y}_{1\bullet})^2} \end{aligned}$$

The first element of LOC_{λ} is thus:

$$\frac{\sqrt{(y_{11}^{\lambda} - \bar{y}_{1\bullet}^{\lambda})^2 + (y_{12}^{\lambda} - \bar{y}_{1\bullet}^{\lambda})^2}}{\bar{y}_{1\bullet}^{\lambda}}$$

$$= \frac{\lambda\sqrt{(y_{11} - \bar{y}_{1\bullet})^2 + (y_{12} - \bar{y}_{1\bullet})^2}}{\lambda \bar{y}_{1\bullet}}$$

$$= \frac{\sqrt{(y_{11} - \bar{y}_{1\bullet})^2 + (y_{12} - \bar{y}_{1\bullet})^2}}{\bar{y}_{1\bullet}}$$

The same is true for each element and thus $LOC_0 = LOC_\lambda$.

F.2 Adding a constant to all outcomes

Consider

		E 1	E 2
I	1	$\tilde{y}_{11}$	$\tilde{y}_{12}$
I	2	$\tilde{y}_{21}$	$\tilde{y}_{22}$

where

$$\tilde{y}_{jk} = y_{jk} + \lambda$$

F.2.1 Adding a constant to all four outcomes does not change the expected return

$$\begin{aligned} E[\tilde{R}] &= P_1(\tilde{y}_{21} - \tilde{y}_{11}) + P_2(\tilde{y}_{22} - \tilde{y}_{12}) \\ &= P_1(y_{21} + \lambda - y_{11} - \lambda) + P_2(y_{22} + \lambda - y_{12} - \lambda) \\ &= P_1(y_{21} - y_{11}) + P_2(y_{22} - y_{12}) = E[R_0] \end{aligned}$$

F.2.2 In nearly all cases, adding a constant to all outcomes changes locus of control

There are four elements in $\tilde{LOC}$. From left to right and top to bottom, call these four elements $\tilde{N}_{1\bullet}$, $\tilde{N}_{2\bullet}$, $\tilde{N}_{\bullet 1}$, and $\tilde{N}_{\bullet 2}$. The numerator and denominator of each element is of the same form. Consider the numerator of the first element.

Consider the element $\tilde{N}_{1\bullet}$ in $\tilde{LOC}$. First, the denominator:

$$\begin{aligned} \bar{\tilde{y}}_{1\bullet} &= \frac{1}{2}(\tilde{y}_{11} + \tilde{y}_{12}) = \frac{1}{2}(y_{11} + \lambda + y_{12} + \lambda) \\ &= \frac{1}{2}(y_{11} + y_{12} + 2\lambda) = \frac{1}{2}(y_{11} + y_{12}) + \lambda \\ &= \bar{y}_{1\bullet} + \lambda \end{aligned}$$

The fact that λ is a positive constant means that I can rewrite the denominator as: The denominator of each element can be written as

$$\bar{y}_{1\bullet} + \lambda = \mu_{1\bullet}\bar{y}_{1\bullet} \tag{22}$$

where

$$\mu_{1\bullet} = \frac{\bar{y}_{1\bullet} + \lambda}{\bar{y}_{1\bullet}} \tag{23}$$

Now, I show that the numerator does not change with the addition of a positive constant:

$$\begin{aligned} &\sqrt{(\tilde{y}_{11} - \bar{\tilde{y}}_{1\bullet})^2 + (\tilde{y}_{12} - \bar{\tilde{y}}_{1\bullet})^2} \\ =&\sqrt{(y_{11} + \lambda + \bar{y}_{1\bullet} - \lambda)^2 + (y_{12} + \lambda - \bar{y}_{1\bullet} - \lambda)^2} \\ =&\sqrt{(y_{11} - \bar{y}_{1\bullet})^2 + (y_{12} - \bar{y}_{1\bullet})^2} \end{aligned}$$

I can then write element $\tilde{N}_{1\bullet}$ as:

$$\tilde{N}_{1\bullet} = \frac{1}{\mu_{1\bullet}} \cdot N_{1\bullet}$$

The same is true for the other elements. $L\tilde{O}C$ can then be written:

$$L\tilde{O}C = \frac{\frac{1}{\mu_{1\bullet}} \cdot N_{1\bullet} + \frac{1}{\mu_{2\bullet}} \cdot N_{2\bullet}}{\frac{1}{\mu_{\bullet 1}} \cdot N_{\bullet 1} + \frac{1}{\mu_{\bullet 2}} \cdot N_{\bullet 2}}$$

Thus $L\tilde{O}C = LOC_0$ only if:

$$\frac{1}{\mu_{1\bullet}} \cdot N_{1\bullet} + \frac{1}{\mu_{2\bullet}} \cdot N_{2\bullet} = \gamma \cdot (N_{1\bullet} + N_{2\bullet})$$

And

$$\frac{1}{\mu_{\bullet 1}} \cdot N_{\bullet 1} + \frac{1}{\mu_{\bullet 2}} \cdot N_{\bullet 2} = \gamma \cdot (N_{\bullet 1} + N_{\bullet 2})$$

Where γ is a positive constant. This perfect balance in proportionality seems unlikely and, indeed, in numerical simulations it proves to be so.

G Examples of External and Internal Maize-specific Locus of Control

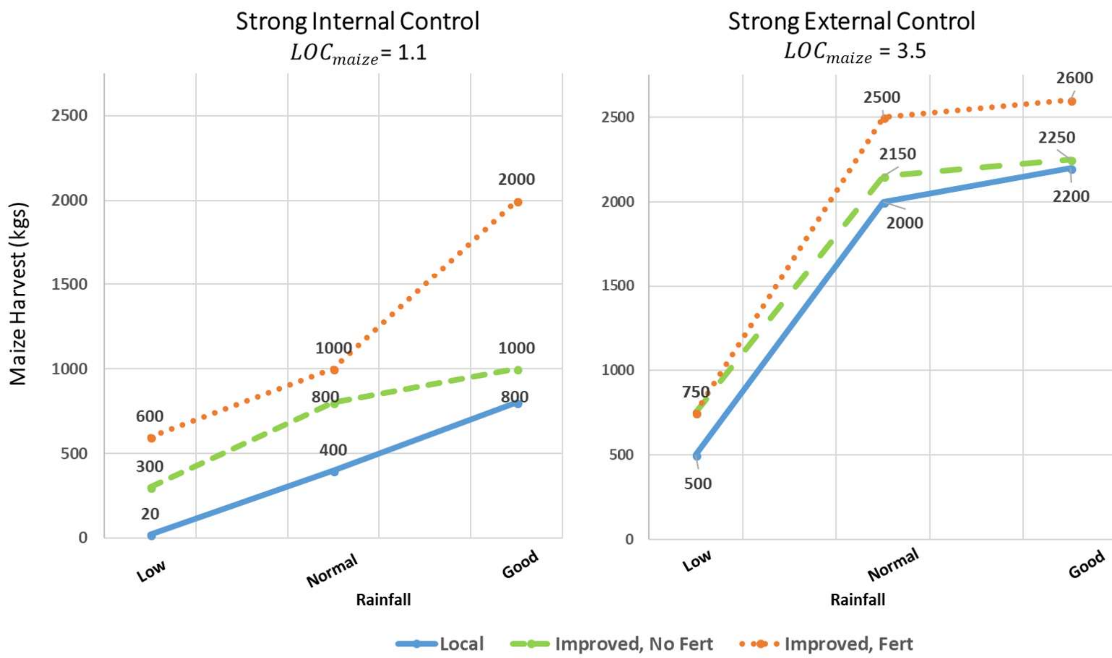

Figure A.8: Hypothetical Harvest Activity - Empirical Examples

H Conditioning on Baseline Beliefs and Baseline Improved Seed Use

Table A.3: Pooled OLS - Effect of Locus of Control on Use of Improved Maize Seed

	Current (1)	(2)	(3)
Maize-specific LOC			
Tercile 2	-0.0841***	-0.0662***	-0.0839***
	(0.0164)	(0.0144)	(0.0172)
Tercile 3	-0.112***	-0.0827***	-0.114***
	(0.0187)	(0.0182)	(0.0198)
General LOC			
Tercile 2	-0.0543***	-0.0365**	-0.0534**
	(0.0197)	(0.0179)	(0.0207)
Tercile 3	-0.0857***	-0.0586***	-0.0805***
	(0.0223)	(0.0210)	(0.0231)
Control for Baseline Difference			
Used Improved Seed		0.301***	
		(0.0283)	
Baseline Expected Return			0.0000144
			(0.00000976)
Add. Controls	No	No	No

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: All specifications include indicator variables for country and year.

Table A.4: Random Effects - Effect of Locus of Control on Use of Improved Maize Seed

	Current (1)	(2)	(3)
Maize-specific LOC			
Tercile 2	-0.0556***	-0.0459***	-0.0559***
	(0.0158)	(0.0155)	(0.0163)
Tercile 3	-0.0745***	-0.0579***	-0.0759***
	(0.0169)	(0.0166)	(0.0175)
General LOC			
Tercile 2	-0.0232	-0.0134	-0.0238
	(0.0170)	(0.0166)	(0.0174)
Tercile 3	-0.0344*	-0.0200	-0.0303
	(0.0188)	(0.0184)	(0.0193)
Control for Baseline Difference			
Used Improved Seed		0.298***	
		(0.0170)	
Baseline Expected Return			0.0000114
			(0.00000838)
Add. Controls	Yes	Yes	Yes
Fixed Effects	No	No	No

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: All specifications include indicator variables for country and year. The Additional Control set includes the respondent household's simple poverty scorecard score, number of maize plots managed, education level of the household head, risk index score, subjective rainfall probabilities, and an indicator for having experienced drought during the previous maize growing season.

Table A.5: Random Effects - Effect of Locus of Control on Use of Improved Maize Seed

	Current (1)	(2)	(3)
Maize-specific LOC			
Tercile 2	-0.0237	-0.0221	-0.0238
	(0.0149)	(0.0149)	(0.0154)
Tercile 3	-0.0300*	-0.0268*	-0.0306*
	(0.0161)	(0.0160)	(0.0166)
General LOC			
Tercile 2	-0.00993	-0.00707	-0.00710
	(0.0159)	(0.0159)	(0.0163)
Tercile 3	-0.0120	-0.00850	-0.00608
Control for Baseline Difference			
Used Improved Seed		0.103***	
		(0.0199)	
Baseline Expected Return			0.0000174**
			(0.00000756)
Add. Controls	Yes	Yes	Yes
Fixed Effects	Village	Village	Village

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: All specifications include indicator variables for country and year. The Additional Control set includes the respondent household's simple poverty scorecard score, number of maize plots managed, education level of the household head, risk index score, subjective rainfall probabilities, and an indicator for having experienced drought during the previous maize growing season.