FILLING A NICHE? THE MAIZE PRODUCTIVITY IMPACTS OF ADAPTIVE BREEDING BY A LOCAL SEED COMPANY IN KENYA

Samuel S. Bird
Michael R. Carter
Travis J. Lybbert
Mary Mathenge
Timothy Njagi
Emilia Tjernström

Working Paper 27636
http://www.nber.org/papers/w27636

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2020, Revised November 2020

This study was commissioned by Acumen, a non-profit impact investment firm, and made possible through the generous support of Acumen and the American people through the United States Agency for International Development Cooperative Agreement No. AID-OAA-L-12-00001 with the BASIS Feed the Future Innovation Lab. We also thank the Agricultural Technology Adoption Initiative (ATAI) administered by JPAL at MIT and the Bill and Melinda Gates Foundation for funding. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Filling a Niche? The Maize Productivity Impacts of Adaptive Breeding by a Local Seed Company in Kenya
Samuel S. Bird, Michael R. Carter, Travis J. Lybbert, Mary Mathenge, Timothy Njagi, and Emilia Tjernström
NBER Working Paper No. 27636
August 2020, Revised November 2020
JEL No. O13,O36,Q12,Q16

ABSTRACT

This paper explores the idea that competitive seed systems may underserve farmers in small, agro-ecological niches, leaving those farmers less productive and poorer than they need be. We develop a theoretical model of the confluence of demand and supply factors that can result in such an equilibrium. We then empirically study the disruption of the maize seed market in Western Kenya that took place when public sector foundation breeding and social impact investment capital together allowed a local seed company to expand and target the area with adaptively-bred maize varieties. A three-year RCT reveals that these seed varieties increased farmer yields and revenues, both for better-resourced farmers (who used non-adapted hybrids and fertilizer prior to the intervention) as well less well-resourced farmers (who did not). This theoretical and empirical evidence suggests news ways for thinking about seeds systems in areas typified by high levels of agro-ecological heterogeneity.

Samuel S. Bird
Department of Economics,
Accounting and Management
Luther College
Decorah, IA 52101
birdsa01@luther.edu

Michael R. Carter
Department of Agricultural and Resource Economics
University of California, Davis
One Shields Avenue
Davis, CA 95616
and NBER
mrcarter@ucdavis.edu

Travis J. Lybbert
Agricultural and Resources Economics
University of California, Davis
1 Shields Avenue
Davis, CA 95616
tlybbert@ucdavis.edu

Mary Mathenge
Department of Agricultural Economics
and Agribusiness Management
Egerton University
Njoro, Naku
Kenya
mathengemk@gmail.com

Timothy Njagi
Tegemeo Institute of Agricultural
Policy and Development
Egerton University
Kenya
tnjagi@tegemeo.org

Emilia Tjernström
School of Economics
University of Sydney
A02 - Social Sciences Building
NSW 2006
Australia
emilia.tjernstrom@sydney.edu.au

1 Introduction

Since 1960, agricultural productivity in sub-Saharan Africa has grown much more slowly than in other regions in the developing world. In 1960, cereal yields in sub-Saharan Africa were just below those of yields in Asia and Latin America. By the early 1990s, this yield gap had more than tripled to 1.5 tons per-hectare, and by 2017, the gap had doubled again to over 3 tons per-hectare (Carter *et al.*, forthcoming). The gap is largely attributable to the region's failure to adopt improved green revolution cereal varieties and other complementary inputs (Evenson & Gollin, 2003), with fewer than half of Sub-Saharan African farmers employing improved varieties, in contrast to near universal adoption elsewhere. The puzzle of this persistently low adoption rate has motivated a large literature, which has identified a range of constraints, ranging from behavioral biases and other internal or psychological constraints (Abay *et al.*, 2017; Duflo *et al.*, 2011), to information (Carter *et al.*, forthcoming), risk (Dercon & Christiaensen, 2011; Karlan *et al.*, 2014), and biophysical and other external resource constraints (Marenya & Barrett, 2009; Suri, 2011). This paper explores an alternative, and ultimately complementary, explanation for this puzzle, highlighting supply-side constraints to the innovation of green revolution varieties appropriate for adoption in sub-Saharan Africa.

In many ways, Kenya is in an exception to the sub-Saharan African pattern of low adoption of improved cereal varieties. By the mid-1980s, a majority of Kenyan farmers had in fact adopted hybrid maize varieties. However, this high average rate of adoption obscures important heterogeneity across regions of Kenya. As shown in Figure 1, in Kenya's two largest maize growing regions, the highland and transitional zones, adoption rates have been above 75% since the 1970s. These hybrid adoption rates contrast with the mid-altitude zone, where this rate has barely crept above 25% in the years since 1970.[1] Foreshadowing

---

[1] Throughout this paper we use the agro-ecological zones for maize production in Kenya developed by the International Center for Improvement of Maize and Wheat (CIMMYT) (Hassan, 1998). These zones are defined by length of the maize growing season, which itself depends in large part on temperature (degree days), as well as rainfall and altitude. This paper focuses on the main zones for maize production: the highland tropical (henceforth, highland) zone composed of farms at or above 1600-2900 meters in altitude,

later discussion, note that the mid-altitude zone is relatively small, constituting only 11% of Kenya's maize area, and yet is home to substantial numbers of Kenya's poor farm households (Hassan, 1998).

This regional divergence in hybrid maize adoption raises the question of its root causes. Regional differences in adoption are not unique to Kenya. In his classic work on technological change, Griliches (1957; 1960) studied the uneven rate of hybrid maize adoption across the US over the decades stretching from the 1930s to the 1950s. He attributed differential adoption patterns to the uneven rate at which locally adapted hybrid varieties[2] became available based on regional differences in market size, public sector investment, and other similar factors. Writing about Kenya, Gerhart (1975) hypothesized that hybrid varieties may have been unprofitable to adopt in some regions because the varieties themselves were poorly adapted to local agro-ecological conditions. If Gerhart's hypothesis is correct, then we would expect the introduction of well-adapted varieties to lead to increased adoption and agricultural productivity. At the same time, even if it is correct, this explanation begs the question as to why well-adapted varieties historically have not been available in this zone.

This paper studies the introduction of seed varieties by a local firm dedicated to adaptive breeding for the agro-ecology of Kenya's mid-altitude zone. To identify the impact of this firm's varieties, we take advantage of a unique opportunity to conduct a randomized controlled trial (RCT) of seed varieties among farmers in communities of western Kenya. Additionally, we provide theoretically-grounded insight, consistent with the reality of the experiment, on the question as to why prior seed sector actors had failed to market locally well-adapted varieties. The results from both exercises all us to draw out implications for the roles of private and public sectors in the seed market.

Our theoretical analysis begins with a model of farmer choice between three stylized

---

the moist transitional (henceforth, transitional) zone from 1200-2000 meters in altitude, and the moist mid-altitude (henceforth, mid-altitude) zone from 1110-1500 meters in altitude. Figure 1 is constructed based on the best available data from these zones, which came from Gerhart (1975), Hassan (1998), TAMPA2 (2004), TAPRA (2010), and the baseline survey for our study.

[2] Adaptive breeding is the process of tailoring improved varieties to specic agro-ecological niches in order to address locally-specific problems or environments.

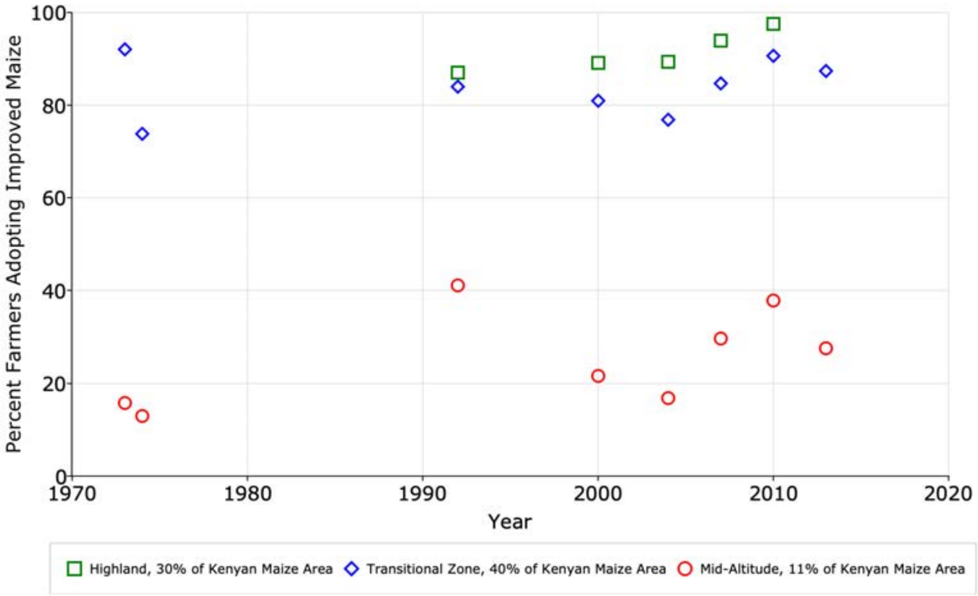


Figure 1: Persistently low adoption of improved maize in Kenya's mid-altitude zone.

seed varieties: local seeds, non-adapted improved seeds, and locally adapted improved seeds. When the only varieties available are local and non-adapted improved seeds, the model predicts that the use of improved seed follows a bi-modal pattern seen in Figure 1, with poorer farmers using only local seeds and better-off farmers using both non-adapted seeds and complementary inputs like fertilizer. The model also shows that the introduction of an improved variety that is well adapted to the local agro-climate can break this pattern and generate benefits to less well-off farmers, as well as to those better-off farmers who had previously adopted improved seeds and inorganic fertilizer.

If farmers can benefit from locally adapted adapted varieties, then why are such varieties not already available in small, "niche" agro-ecologies like Kenya's mid-altitude zone? Using a simple model of seed variety innovation, we show that absent public-private partnerships in the seed sector, and absent buoyant access to capital for locally-based seed companies, the equilibrium outcome in niche agro-ecologies is one of no innovation of locally adapted varieties, leading to low adoption and low productivity. However the model suggests that availability of public-private partnerships and capital to locally-based seed companies may spur the seed innovation needed to break this trap.

To test the empirical veracity of this theoretically-grounded story, we take advantage of a unique setting created by the expansion in Kenya of a local seed company, Western Seed Company (henceforth, Western Seed). Since the early 2000s, Western Seed has benefited from strong public-private partnership with the International Center for Improvement of Maize and Wheat (CIMMYT) to innovate well-adapted improved varieties for the mid-altitude zone. But for much of that time Western Seed's market coverage in Kenya was limited due to capital constraints, a primary bottleneck for seed companies in the region (Langyintuo *et al.*, 2010). In recent years, Western Seed has benefited from a partnership with the social impact investor Acumen to increase the capital to fund its market expansion. Western Seed's expansion allowed our research team to conduct a RCT to study the impact of their locally adapted, improved maize varieties.

In line with the expectations of the theoretical model, the RCT reveals that the introduction of Western Seed varieties caused substantial yield gains for maize farmers in the mid-altitude zone. Outside of this zone, we find that the new seed varieties provided by Western Seed performed no differently than the improved varieties already available in the market. Also as predicted by the theoretical model, the effects in the mid-altitude zone vary across farmers. For farmers who historically did not use improved maize varieties, our intention to treat impact estimates reveal that the availability of Western Seed varieties increased yields by 21% on average, a large effect despite these farmers using little to no complementary inputs like fertilizers.[3] However we would expect farmers who historically used improved seeds to have done so in part because they had the resources to invest in complementary inputs like fertilizers. We find this to be true empirically, and that these better-resourced farmers realized an even average yield gain of 47% percent due to the availability of Western Seed varieties.

Stepping back from Kenya, sub-Saharan Africa is known to be comprised of a wide-

---

[3] With compliance rates just south of 20% for this group, impacts on those who actually adapted the seeds are substantially higher, as discussed later in the paper.

variety of different agro-ecologies. We cannot in this paper pin down the extent to which the broader sub-Saharan African pattern of low use of green revolution technologies can be attributed to niche agro-ecologies for which locally adapted improved varieties do not exist. But our results point to a public-private seed sector model that potentially offers benefits across the wealth spectrum of African cereal farmers.

The remainder of this paper is structured as follows. Section 2 models both the demand for and supply of improved seed varieties in a small agro-ecological niche. Section 3 introduces the western Kenya study area and lays out the design for the RCT made possible by the capacity expansion of Western Seed. Section 4 presents average treatment effects for the mid-altitude zone as well as for other zones included in our study. Section 5 undertakes the key heterogeneity analysis, identifying the impacts within the mid-altitude zones on farmers that had and had not been prior users of hybrid seeds. The final section concludes with reflections on implications for seed systems.

2 Economics of Technological Change in a Niche Agro-Ecology

This section models the demand and supply for three stylized seed varieties in an agro-ecological niche market, such as Kenya's mid-altitude zone. The varieties differ in their average yields, their substitutability across environments, and their responsiveness to fertilizer. A local variety is retained from the previous season's harvest and is relatively well-adapted to the local environment but unresponsive to complementary fertilizer applications. The second variety is improved and fertilizer-responsive, but was developed for a different agro-ecological zone without further adaptation to the local environment. We refer to this variety as the non-locally-adapted (NLA) improved variety. Finally, there is an improved variety that is fertilizer-responsive and has been adaptively bred for the local agro-ecological conditions.

The production environment is characterized by linear yield responses to changes in fertilizer application. This is a reasonable specification over the relevant range of small farmer fertilizer use if maize plants exhibit increasing returns to nutrition at low nutrient levels and if farmers optimally manage fertilizer application rates (see Appendix A below) . We thus write per-hectare yields as a function of seed type and fertilizer:

$$y^v(f^v) = (\alpha_0^v + \alpha_1^v f^v) \quad \forall f < f^o,$$

where the variety indicator v takes on the value of r for retained local seed, n for NLA improved variety and a for locally-adapted improved variety. The term f^v is the per-hectare intensity of fertilizer applied to variety v and f^o is the agronomically optimal fertilizer rate.[4]

Assigning a numeraire price of one for maize, a farmer who devotes H^v hectares of land to variety v will earn the following net income (i.e., less input costs):

$$Y^v = H^v \left((\alpha_0^v - s^v) + (\alpha_1^v - p_f)\, f^v \right)$$

where s^v is the per-hectare cost of seed for variety type v, and p_f is the price of fertilizer.

Using this notation, we characterize the three stylized seed technologies as follows:

- *Retained local variety* (r)

  The per-hectare cost of local seeds, s^r, is low (since farmers can save grain from the previous harvest). Fertilizer application is not profitable ($p_f > \alpha_1^r$).

- *Non-locally adapted, improved variety* (n)

  The cost of NLA seeds is higher than local seeds ($s^n > s^r$), and while this variety is fertilizer-responsive ($\alpha_1^n > p_f$), its unsuitability to the local environment means that it is less profitable than local varieties withou complementary fertilizer application, i.e.

---

[4] In reality, returns to fertilizer, α_1^v, vary with soil quality (Tjernstrom 2017), farmer skill (Laajaj & Macours, forthcoming) and weather. Adding in these additional dimensions of farm or farmer heterogeneity would add complexity (e.g., Figure 3 would increase in dimensionality and the cutoffs would become frontiers) but little additional insight. As we discuss in more depth below, incorporating weather or other stochastic factors that affect input productivity would likely strengthen the qualitative characteristic of the primary results.

Table 1: Benefits from introducing an adapted hybrid in a niche agro-ecological zone.

	Adapted Hybrid More Fertilizer Responsive ($\alpha_1^a > \alpha_1^n$)	**Non-Adapted Hybrid More Fertilizer Responsive** ($\alpha_1^a < \alpha_1^n$)
Adapted Hybrid Outperforms Retained Variety without Fertilizer ($\alpha_0^a - s^a > \alpha_0^r - s^r$)	*Case 1* (Both Groups Benefit)	*Case 2* (Only Less Wealthy Benefit)
Retained Variety Outperforms Adapted Hybrid without Fertilizer ($\alpha_0^a - s^a < \alpha_0^r - s^r$)	*Case 3* (Only Wealthier Benefit)	*Case 4* (Neither Group Benefits)

$((\alpha_0^n - s^n) < (\alpha_0^r - s^r))$.

- *Locally-adapted improved variety* (a)

  Also more costly to the farmer than retained seeds, we assume that these varieties are no more costly than the NLA hybrid ($s^n \geq s^a > s^r$).

The key question this study addresses is whether a locally adapted hybrid can economically outperform the retained and NLA seed alternatives. Table 1 shows the different possible cases. In Case 1, the locally adapted variety would outperform the retained variety even when fertilizer is not used, and it would also outperform the NLA variety when fertilizer is used. Case 4 is the opposite case, with the locally adapted variety outperforming neither retained nor NLA varieties under these conditions. The off-diagonal cases (2 and 3) are where the the locally adapted variety outperforms one type of seed-fertilizer combination, but not the other. Later empirical analysis from the RCT will allow us to test these assumptions in the case of varieties bred by one seed firm for Kenya's mid-altitude zone.

Figure 2 portrays these different technological options using the stylized representations of retained seeds and NLA hybrids, and Cases 1 and 4 for the locally adapted hybrid. On the x-axis, $\tilde{f}^n$ indicates the fertilizer level below which the NLA is less profitable than the local variety. Fertilizer use is wasted on the local variety since its yields do not increase with fertilizer. In Case 1, the locally adapted improved variety outperforms the alternatives irrespective of the level of fertilizer use. In Case 4, it outperforms neither.

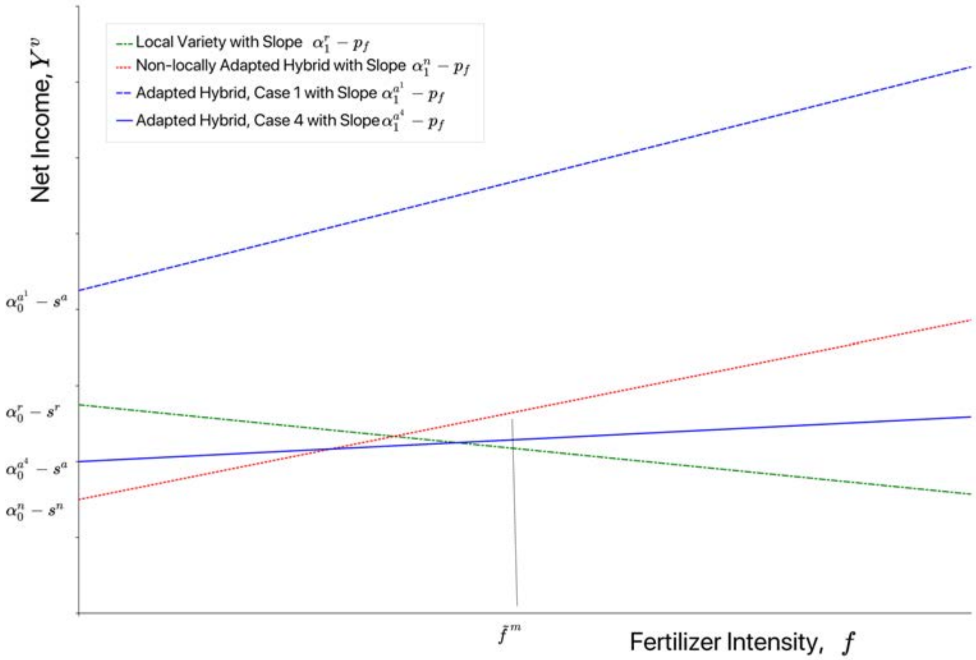


Figure 2: Income under different seed varieties.

We build on our stylized representation of alternative seed technologies to explore the demand and supply of improved technologies in a niche agro-ecology. Since an empirical test of the efficacy of adaptive breeding in general is not feasible, our ultimate goal in this paper is to empirically test the efficacy of specific adaptively-bred varieties using data from our RCT. If these particular adaptively-bred varieties do not lead to farm-level impacts, we cannot conclude that that adaptive breeding in general is ineffective, as it may be that the particular breeder we examine was ineffective. However evidence of impacts of these adaptively-bred varieties would support the efficacy of adaptive breeding.

2.1 Demand for Improved Seeds under Liquidity Constraints

While improved seed varieties may have the potential to increase the incomes of farmers, farmer demand for improved seeds faces several constraints. In our model, farmers face two constraints. First, the farm household lacks access to credit markets and hence must self-finance input purchases using prior earnings, or what we will call cash-on-hand, z_1. Input purchases thus compete directly with current consumption needs.

Second, improved seed varieties are packaged in bags of no less than two kilograms in Kenya (and elsewhere), which is enough to sow one-fifth of an acre. This limitation creates a minimum farm area that can be devoted to improved seed varieties, which we denote $\underline{\mathrm{H}}^v$, with $\underline{\mathrm{H}}^r = 0$ and $\underline{\mathrm{H}}^n, \underline{\mathrm{H}}^a > 0$. While seed bags can in principal be opened and sub-divided, we assume that the well-known problem of counterfeit seeds makes farmers unwilling to buy less than a single bag of seed.[5]

We assume that the farming input choices of a household with $\bar{H}$ units of land and z_1 units of cash-on-hand, are guided by the following two-period model:

$$(1) \quad \underset{\tilde{H}^v, I^v, f^v}{Max} \quad u(c_1) + \beta u(c_2)$$

$$subject\ to:$$

$$c_1 \leq z_1 - \sum_{v=r,n,a} H^v \left(s^v + p^f f^v \right)$$

$$c_2 \leq \sum_{v=r,n,a} H^v \left(\alpha_0^v + \alpha_1^v f^v \right)$$

$$\sum_{v=r,n,a} H^v \leq \bar{H}$$

$$I^v (I^v - 1) = 0, \forall v$$

$$(\tilde{H}^v - \underline{\mathrm{H}}^v) I^v \geq 0, \forall v$$

$$H^v \equiv \tilde{H}^v \cdot I^v, \forall v$$

$$f^v \geq 0, \forall v$$

where $u(\textbf{.})$ is a concave utility and $\beta < 1$ is the per-period discount factor.[6] As written in problem (1), the household chooses the notional area, $\tilde{H}^v$, that it would devote to variety v

---

[5] In our study area, Western Seed and other companies faced such significant problems of counterfeit seeds (unscrupulous individuals would collect used seed bags and refill them with local seeds) that they began enclosing tickets that could be used to certify the seeds' authenticity through an SMS-based message system.

[6] We assume that this per-period discount factor is greater than the rate of return on savings so that the household consumes in the first period rather than saving for the second period.

if there were no seed bag size limit. The constraints then transform this notional area to an effectively cultivated area of 0 if $\tilde{H}^v < \underline{\mathrm{H}}^v$. Otherwise, $H^v = \tilde{H}^v$. Denote the optimum area that will be planted to variety v as $H^{v*}(z_1, \bar{H})$.[7]

For purposes of our analysis here, we are only interested in characterizing when the farm household would begin using improved seed technologies, that is, when it would jump from planting no area to improved varieties to planting at least $\underline{\mathrm{H}}^v$ of those varieties. Given the binary nature of this decision, we only need to compare optimized household well-being when the household only plants the local variety, $V^r(z_1, \bar{H}|H^n, H^a = 0)$, with the stream of utility available when planting $\underline{\mathrm{H}}^v$. We first consider the case when only the NLA improved variety is available, and then consider the implications of introducing a locally adapted improved variety.

2.1.1 Choice of Retained Seeds versus a Non-Locally Adapted Improved Variety

To analyze the farm household's decision to adopt the NLA improved variety, we evaluate optimal behavior per maximization problem (1) conditional on the household devoting the minimum possible area to the NLA improved variety, $\underline{\mathrm{H}}^n$.[8] The first order condition for optimal fertilizer choice under this conditional problem is:

$$\tilde{\lambda}(f^n, z_1)p_f \geq \alpha_1^n$$

where $\tilde{\lambda}(f, z_1) = \frac{u_1'(\tilde{c}_1)}{\beta u_2'(\tilde{c}_2)}$ is a measure of the shadow price of liquidity, $\tilde{c}_1 = z_1$-$\underline{\mathrm{H}}^n\,(s^n + p_f f^n) - \left(\bar{H} - \underline{\mathrm{H}}^n\right) s^r$ and $\tilde{c}_2 = \underline{\mathrm{H}}^n(\alpha_0^n + \alpha_1^n f^n) + \left(\bar{H} - \underline{\mathrm{H}}^n\right) \alpha_0^r$. Note that first period consumption is cash-on-hand less spending on inputs for the NLA variety and the seed costs of the retained variety (recall that under our technology assumptions, fertilizer cannot be profitably applied to the local variety such that $f^r = 0$ under optimal choice).

---

[7] In principal, this seed bag integer problem should continue, but as our analysis only concerns the adoption decision, we will ignore that aspect of the problem.

[8] In what follows, we all assume that it the farm household always finds it optimal to fully cultivate its available land. While this assumption could be relaxed, it would add complexity without additional insight for the problem at hand.

Note that $\tilde{\lambda}(f^n, z_1)$ is increasing in fertilizer investment and decreasing in cash-on-hand. An interior solution with $f^n > 0$ would equate the marginal returns to fertilizer to its full cost, marked up by the shadow price of liquidity. For a sufficiently poor household with low z_1, the full cost of fertilizer will always exceed its return even when evaluated at $f^n = 0$. Conditional on adopting the improved variety, the best such a household could do would be to apply no fertilizer. However, under the technology assumptions above, we know that the household will generate less spendable income employing the NLA variety any time it chooses $f^n < \tilde{f}^n$. As can be more formally demonstrated, for these sufficiently poor households, optimized utility conditional on minimal adoption of the NLA improved variety, $V^e(z_1, \bar{H}|H^n = \underline{\mathrm{H}}^n)$, will be strictly less than the non-adoption alternative, $V^\ell(z_1, \bar{H}|H^n, H^a = 0)$. Note that a more realistic model that included risk considerations would only serve to strengthen this result.[9]

More generally, defining the difference between these two value functions as

$$\Delta^{nr} = V^n - V^r$$

we can show that $\frac{\partial \Delta^{nr}}{\partial z_1} > 0$ and that there is a critical cash on hand level, $\tilde{z}^n$ where Δ^{nr} becomes positive and adoption of the improved NLA variety will begin with fertilizer levels no less than $\tilde{f}^n$ (see Appendix B).

Letting $\phi(z_1)$ and $\Phi(z_1)$ respectively denote the *pdf* and *cdf* for the distribution of initial cash-on-hand in the small agro-ecological niche, Figure 3 illustrates the implications of this model for the demand for NLA improved seeds. Only farm households with initial cash in excess of $\tilde{z}^{nf}$ would demand the NLA improved seed variety and optimally apply a non-trivial amount of fertilizer to them. In short, when only the NLA improved and retained varieties are available, we would expect to see bifurcation, with $1 - \Phi(\tilde{z}^{nf})$ fraction of the population

---

[9] In a model with risk, fertilizer applied to the NLA variety would increase expected yields as well as the variance of net income under the defensible assumption that fertilizers do not decrease yield variance. Assuming households are risk averse, they would apply even less fertilizer under risk than they would in our simplified model with certain yields. This would increase the benefits from planting the retained variety without fertilizer over the NLA variety with low fertilizer use relative to the case without risk.

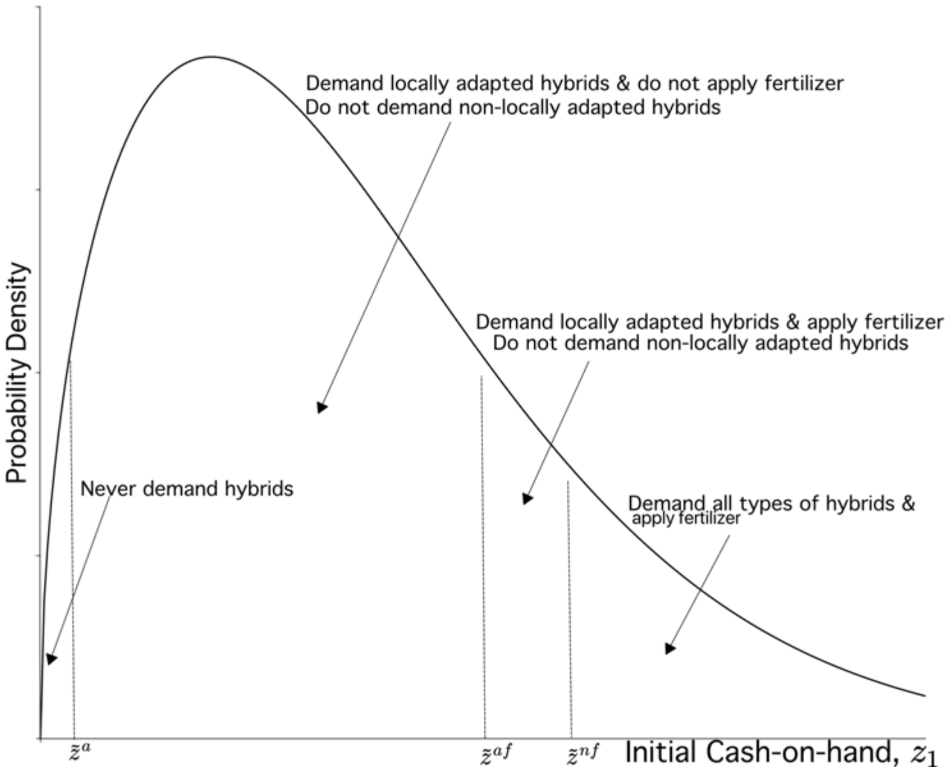


Diagram assumes the Case 1 parameter values in which adaptive breeding works.

Figure 3: Initial cash-on-hand distribution and thresholds for adoption of improved seeds.

adopting (and applying fertilizer) to the improved varieties, and the complementary fraction ($\Phi(\tilde{z}^{nf})$) employing only local, retained varieties without fertilizers. The magnitude of these fractions of course depends on the actual distribution of cash-on-hand and the magnitude of the key technology parameters.

2.1.2 Demand for Locally Adapted Improved Varieties

We now consider how farm household variety choice changes after the introduction of a locally adapted improved seed variety. Under Case 1 in Table 1 (where adaptive breeding offers gains both to famers who can afford fertilizers as well as to those who cannot), no farm household would ever adopt the NLA variety rather than the adapted variety. There will again appear a critical cash-on-hand value, $\tilde{z}^{a}$, at which the farm household will find it optimal to adopt the minimum acreage to the locally adapted improved variety. In contrast to the NLA variety, some farmers who cannot afford to purchase fertilizer will adopt the locally adapted improved variety (under the Case 1 assumption that $(\alpha_0^a - s^a) > (\alpha_0^r - s^r)$). There will appear a second critical level of cash-on-hand, $\tilde{z}^{af} < \tilde{z}^{nf}$, at which the farm household will adopt the improved variety and apply a positive amount of fertilizer (see Appendix B for more details).

The implications of this Case 1 analysis are again illustrated in Figure 3. With the introduction of the locally adapted improved variety, the farming population now falls into three sub-groups:

1. $\Phi(\tilde{z}^a)$ of the population will continue to cultivate only retained, non-improved varieties;
2. $\Phi(\tilde{z}^{af}) - \Phi(\tilde{z}^a)$ of the population will cultivate the locally adapted improved varieties without using fertilizers, but will enjoy higher yields than a counterfactual group that had access to only NLA varieties and chose neither use them nor employ fertilizer; and,
3. $1 - \Phi(\tilde{z}^{af})$ of the population will deploy the new varieties and apply fertilizer. Note that this group would be predicted to have higher yields than a counterfactual group that only had access to NLA varieties and utilized them in conjunction with fertilizer.

All of these implications of course depend critically on the Case 1 parameter assumptions about the relative productivity of locally adapted hybrids. Under the Case 4 assumptions, the introduction of the locally adapted varieties would be meaningless as no farmer would adopt them. Cases 2 and 3 are the intermediate cases. In Case 3, some farmers would switch from retained varieties to the improved varieties but better endowed farmers would stay with the NLA hybrid (i.e., $\tilde{z}^a$ would exist, but $\tilde{z}^{af}$ would not). In Case 4, we would have the opposite configuration with less well-endowed farmers sticking strictly with retained varieties and the better off switching to the locally adapted hybrid.

After examining the circumstances under which the seed sector will supply locally adapted varieties, we will return to the experience of Western Seed in western Kenya to see if, in fact, local adaptive breeding can provide benefits to population sub-groups 2 and 3. In the specific case of the Western Seed hybrids studied below, the company itself advertised that farmers who moved from retained seeds without fertilizer to Western Seed hybrids without fertilizer would experience a 50 percent yield increase. The company further suggested that proper fertilization could boost yield by 300 percent compared to farmer-retained seeds (Partners, no date). A separate document prepared by an impact investment fund suggested that Western Seed hybrids would offer a 25 percent yield increase compared to other hybrid maize varieties (Fund, 2010). In short, these promotion documents claim that Case 1 is correct, a claim that we will later examine empirically using data from our RCT.

2.2 Supply of Locally Adapted Improved Seed Varieties

If farmers are willing to adopt locally adapted seeds (as some would be true in all but Case 4 above), then this raises the question of what conditions are needed to support private seed companies to introduce adapted seed varieties in a small, niche agro-ecology. In other words, what accounts for the 30 years of a stagnant hybrid adoption rate in Kenya's mid-altitude zone? One explanation is that improved varieties cannot be successfully adapted to this region (Case 4 in Table 1, above). In this section, we consider an alternative explanation by putting forward a stylized model concerning the innovation and supply of improved, locally adapted seed varieties. Innovate here means to conduct a breeding program that leads to the development and marketing of a locally adapted improved variety. Our distinction between innovation and supply of seed varieties is analogous to the innovation and product markets in the conceptual model of Spielman *et al.* (2014).

In our model, a key factor is how the firm obtains parent lines for local adaptive breeding. If parent lines are owned and maintained by the firm itself, then the firm pays no royalties for using the parent varieties. Given the high cost of developing and maintaining parent breeding lines, this option is only available to large, multinational firms. Smaller firms can access parent material by purchasing use rights from other private sector firms at a royalty cost of ρ per-kilogram of adapted seed produced. In addition, firms may have the option to use without royalty parent seed produced by the public sector.[10] When public sector breeders make parent lines freely available, smaller firms can avoid paying royalty costs for parent material.

Given access to parent seed lines, firms incur a non-trivial fixed cost to breed seed varieties for the local agro-ecology. These costs are related to not only acquiring farm land on which to experiment, but also to acquiring the knowledge about local conditions that limit crop performance. We assume that local firms, which typically emerge from farms already producing in the local area, have a fixed cost advantage over multinational or other non-local firms who need to both acquire land and learn about the particularities of local farm production. Specifically we assume that $F^\ell < F^m$, where the first term measures the fixed costs of adaptive breeding for the local firm and the second the fixed costs for the multinational or non-local firm.

Taking seed prices as given,[11] we can write the short-term returns to innovation for three types of seed companies:

$$\begin{aligned}\pi^\ell(q) &= q\left(p^a - c - \rho\right) - F^\ell \\ \pi^{\ell p}(q) &= q\left(p^a - c\right) - F^\ell \\ \pi^m(q) &= q\left(p^a - c\right) - F^m\end{aligned}$$

where q is the quantity of adapted seed sold, p^a is its market price, and c is the per-kilogram production or seed multiplication costs once the variety mix is established. The term $\pi^\ell(q)$ measures returns for a local firm that must pay royalties for parent seed, $\pi^{\ell p}(q)$ are returns for a local company with a public partnership that allows it to freely access quality parent seed, and $\pi^m(q)$ are returns for the multinational or non-local firm. Figure 4 graphs these

---

[10] The public sector has long filled this role, which has, for example, contributed to regional differences in the development and adoption of hybrid maize in the United States (Griliches, 1960; Kantor & Whalley, 2019). In developing countries, public sector investments are supplemented by investments by international organizations through the CGIAR networks, in particular CIMMYT for maize.

[11] In this sub-section, we also consider how this problem might change if firms acted as monopolists and set prices for the niche market.

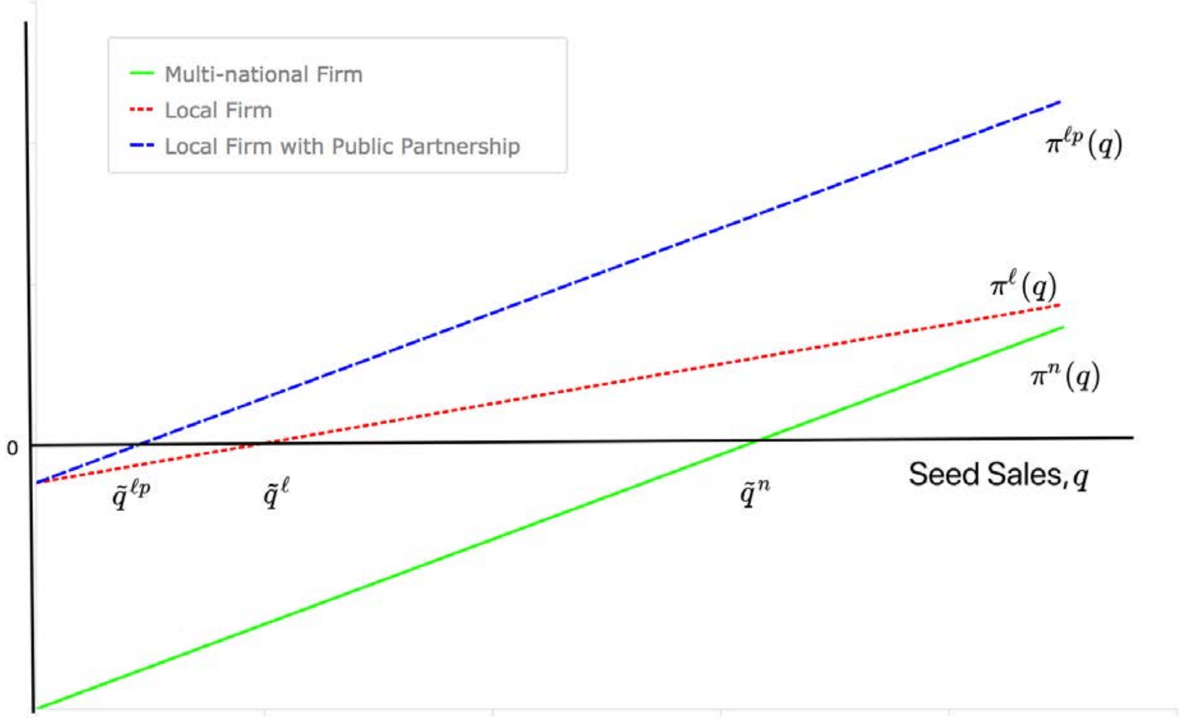


Figure 4: Firm profits for locally adapted seeds as a function of sales.

different functions.

2.2.1 Economics of Innovating Locally Adapted Improved Seed Varieties

To analyze whether a firm chooses to innovate, we evaluate the firm's optimal seed sales conditional on it choosing to innovate. We assume that the marginal variable profits from seed sales are positive for all firm types:

$$p^a > c + x + \rho > c + x$$

Under this assumption, firm profits increase with seed sales conditional on the firm choosing to innovate.[12]

Given the fixed cost structure of local adaptive breeding, we can define for each type of

---

[12] Firms will of course not invest in adapting local varieties if this condition is not met.

firm a critical seed sales volume ($\tilde{q}$) where revenues just cover the fixed and variable costs of production:

$$\tilde{q}^{\ell} \equiv \frac{F^{\ell}}{p^a - c - \rho}, \tilde{q}^{\ell p} \equiv \frac{F^{\ell}}{p^a - c}, \tilde{q}^{m} \equiv \frac{F^{m}}{p^a - c}.$$

For theses sales volumes, firms would recover their fixed costs of innovation, but earn zero profits. Since firm profits increase with seed sales conditional on the firm choosing to innovate, the firm will not innovate for markets smaller than the critical market size.

In what follows, we ignore the entry of other firms offering locally adapted hybrids. Given that our focus is on areas that heretofore have seen no entry, this assumption appears as a reasonable starting point. We could of course imagine more complex structures in which firms have conjectures about the entry of other firms. In this case, the market size constraints we derive below become necessary but not sufficient conditions for innovation of locally adapted varieties and entry into the niche market. If anything, the competitive deterrent of this alternative structure would reenforce the tendency of niche markets remaining underserved by innovation.

Abstracting from potential competition from other seed companies that might engage in locally adaptive breeding, we now focus on two primary constraints that may block innovation and entry into a niche market. The first is the size of the agro-ecological zone for which breeding will be undertaken.[13] Letting N denote the number of farmers in the agro-ecological zone and assuming that each farmer has $\bar{H}$ units of land, then using the notation developed in Section 2.1 above, we can define the effective market demand for adapted improved seeds as the sum of seed demand across all households in the zone given the distribution of wealth or cash-on-hand:

---

[13] In practice, the size constraint is the size of the agro-ecological zone within a nation's borders. Seed industries are highly regulated at the national-level such that seed companies incur fixed costs to entering each individual country. Relaxing these regulatory bottlenecks at the national-level would effectively relax the market size constraint for firms considering investing in adaptive breeding.

$$q^{max} = \frac{s^a}{p^a} N \left(\int_{\tilde{z}^a}^{\tilde{z}^{af}} H^{v*}(z_1)\phi(z_1)dz_1 + \int_{\tilde{z}^{af}}^{\infty} H^{v*}(z_1)\phi(z_1)dz_1 \right)$$

where all the terms are as defined above.[14] As shown above, farm households with cash-on-hand below $\tilde{z}^a$ will not use any of the improved seeds, and hence this group is excluded from the integral terms. Note that this expression highlights two features that shape potential market demand: size of the agro-ecological zone (N) and the distribution of wealth. In practice, we might expect these two objects to be correlated. A too-small zone may not be offered improved adapted varieties and hence more households may be locked into a low productivity, low income trap that keeps many farm households below the critical cash-on-hand levels needed for adoption.[15]

A second constraint that may limit a firm's economic ability to invest in adaptive seed breeding is the firm's own capital constraint. Using the simple notation of our model, we can express this constraint for a local firm without a public partner as:

$$q(c + \rho) + F^{\ell} \leq \bar{K}^{\ell}$$

where $\bar{K}$ is the capital that the firm can leverage for investment. This expression, and its analogue for the other firm types, defines a maximum amount of seed production that the firm can afford to finance, $q_K^{\ell} \equiv \frac{\bar{K} - F^{\ell}}{c+\rho}$. While all firms potentially face capital constraints, capital constraints are most likely bind for local firms (Langyintuo *et al.*, 2010).

Returning to Figure 4, we can now explore the basic intuition from this model. If $q^{max} <$

---

[14] This expression assumes the case 1 parameter values (see Table 1) in which the locally adapted varieties outperforms retained seeds with fertilizer and outperforms non-locally adapted hybrids with fertilizer. Under the other parameter cases, the limits of integration in the above expression would need to be adjusted to eliminate demand from the sub-population(s) of farmers who would not benefit from locally adapted hybrids.

[15] Entering firms could in principal exercise power over the prices they charge. In this case, q^{max} would become a decreasing function of price charged. At the same time, $\tilde{q}$ would decrease with price charged. While it would be possible to write down an integrated monopolist's problem in which seed price is optimally chosen, doing so would not alter the core implications of the model concerning innovation and entry.

$\tilde{q}^m$, the multinational firm will not find it profit-maximizing to invest in adaptive breeding for the local agro-ecological zone. Similarly, even if $q^{max} > \tilde{q}^\ell$, the local firm without a public partner will not invest if $q_K^\ell < \tilde{q}^\ell$. In markets with only these two types of firms, there could be a range of small (niche) agro-ecologies that are simply not supplied with any locally adapted improved seed, resulting in the kind of bifurcated adoption behavior described in Section 2.1 above and illustrated by Figure 1. As we will see in the next section, this appears to have been the case for the relatively small mid-altitude region of Kenya prior to the introduction of locally adapted improved seeds.

2.2.2 Supplying Seed Technologies by Relaxing Capital Constraints

As the prior discussion makes clear, it may not be profitable to invest in the breeding needed to innovate adapted improved seed varieties for niche agro-ecological zones for either large multinational actors or capital-constrained local seed companies. However, as displayed in Figure 4, a well-financed local company with access to publicly provided foundation seed may find it profitable to innovate for those niche zones.

Financing to relax capital constraints also can enable seed companies with locally adapted varieties to expand their production of those varieties. In the case of Kenya described in the introduction, Western Seed not only had a partnership with a public source of quality parent line seed, it also received a major infusion of capital that allowed it to rapidly expand its seed multiplication capacity. As the next sections now explore, the expansion of this seed company gave us the opportunity to explore whether the seed system in Kenya was indeed leaving money on the table by failing to realize profitable innovation of locally adapted improved maize seeds.

3 Empirical Context and Experimental Design

The theoretical model laid out in the prior section demonstrates that while locally adapted improved varieties can benefit both worse- and better-off farmers, there may not be a supply of such varieties to small agro-ecological zones unless capital and public goods are provided to local seed firms that enjoy informational and other cost advantages over multinational and other larger scale competitors. In this section, we first show that Kenya's mid-altitude region largely fits the predictions of the model. We then lay out our research design intended to allow identification of the impact of the introduction of locally adapted improved varieties on the yields of smallholders farmers.

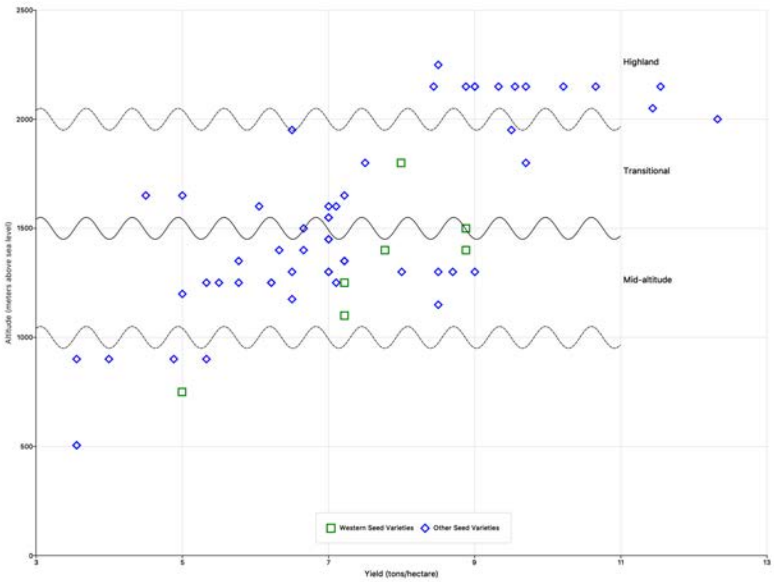
3.1 Maize in Kenya's Mid-altitude Zone

Our stylized model suggests regional differences in adoption seen in Figure 1 may be due to the adaptive breeding of maize hybrids for some zones, but not others. In the 1960s, the government of Kenya supported research and development of hybrid maize for the highland zone, which covers almost one-third of Kenya's maize-growing areas and has high agricultural potential (Gerhart, 1975; Hassan, 1998). Significantly less investment was made to develop varieties for other regions, including the mid-altitude zone. As shown earlier, hybrid varieties were quickly adopted by almost all farmers in the highland zones and the neighboring transitional zone, but hybrid varieties were not adopted elsewhere, including in the mid-altitude zone (Gerhart, 1975).[16]

As shown in Figure 1, these differences in hybrid maize adoption persist to this day. The gap in hybrid maize adoption across regions of Kenya was the focus of an in-depth study to chart a path for future maize research by CIMMYT (Hassan, 1998). The study recommended the development of parent lines for maize varieties that can mature during the shorter growing

---

[16] In essence hybrid maize development in Kenya exemplifies the model from Byerlee & Traxler (2001) of public-sector investment in agricultural research and development. In this model, a budget-constrained country with both small and large agro-ecological zones adopts a two-part strategy with intensive research effort devoted to develop varieties adapted for in the large zone and NLA technologies transferred to the smaller zone.



Source: Commercial recommendations for hybrid varieties from company websites among varieties registered by Kenya Plant health inspectorate Service (KEPHIS, 2018)

Figure 5: Yields and recommended planting altitude for registered maize varieties as of 2010.

seasons typical of the mid-altitude zone of western Kenya. Subsequently, public investment in research and development, as well as reforms to seed markets,[17] spurred private investment in innovation and product markets. In innovation markets, the shift in public-sector research by CIMMYT laid the groundwork for the private sector to develop maize hybrids for the mid-altitude zone of western Kenya. In particular, Western Seed emerged as a leader in varietal research and development for the lower potential, mid-altitude zone of western Kenya.

While the different agro-ecological zones are not defined exclusively based on altitude (see footnote 1 above), the Kenyan Plant Health Inspectorate Service (KEPHIS) issues annual reports on the performance of different seed varieties at a recommended altitude and season length. Using these data, Figure 5 shows that in the 1000-1500 meter altitude range typical of the mid-altitude agro-ecological zone, Western Seed produces many of the varieties with the greatest yields.

---

[17] In the early 1990s, the seed market in Kenya was liberalized.

3.2 Expansion of Western Seed and the Randomized Controlled Trial

With its new varieties, the geographic footprint of Western Seed maize hybrids slowly expanded over time from the transitional zone into the mid-altitude zone near Lake Victoria.[18] However, Western Seed's seed multiplication and market expansion was constrained by its available capital (Partners, no date). In 2008 and in 2010, two impact investment organizations (Pearl Capital Partners and Acumen Fund) made debt and equity investments in Western Seed totaling \$3 million with the intention of rapidly tripling WSC's seed supply capacity. Within a few years, this new supply capacity was on-line, opening the door for both Western Seed's geographic expansion and coincidentally creating a unique opportunity to establish an RCT around the locally adapted hybrids.

In partnership with Western Seed and Acumen, the research team established a research design that would allow identification of the impact of the introduction of WSC hybrids in new areas in western and central Kenya. Specifically, Western Seed had resources to establish 100 new demonstration plots at key points across these regions for the 2013 planting season. Each demonstration plot was designed to provide information to villages within a 5 to 10 mile radius of the plot and the sites were spaced with that distance in mind. At the research team's request, Western Seed identified 125 potential demonstration plot sites (25 more than they wanted) with the understanding that the team would randomly allocate up to 25 sites to a control group where no demonstration plots nor marketing would take place. Matching Western Seed's expansion plan, the sites were spread across western and parts of central Kenya.

Figure 6 maps the study sites across central and western Kenya. The background shading on the mapping shows the different altitude ranges of highland areas (above 2000 meters), transitional areas (1500-2000 meters) and mid-altitude areas (below 1500 meters). While

---

[18] Tegemeo Institute's TAPRA data set (see TAPRA (2010)) allows us to see the expansion of Western Seed hybirds into the mid-altitude zone between the 2004 and 2010 TAPRA survey rounds.

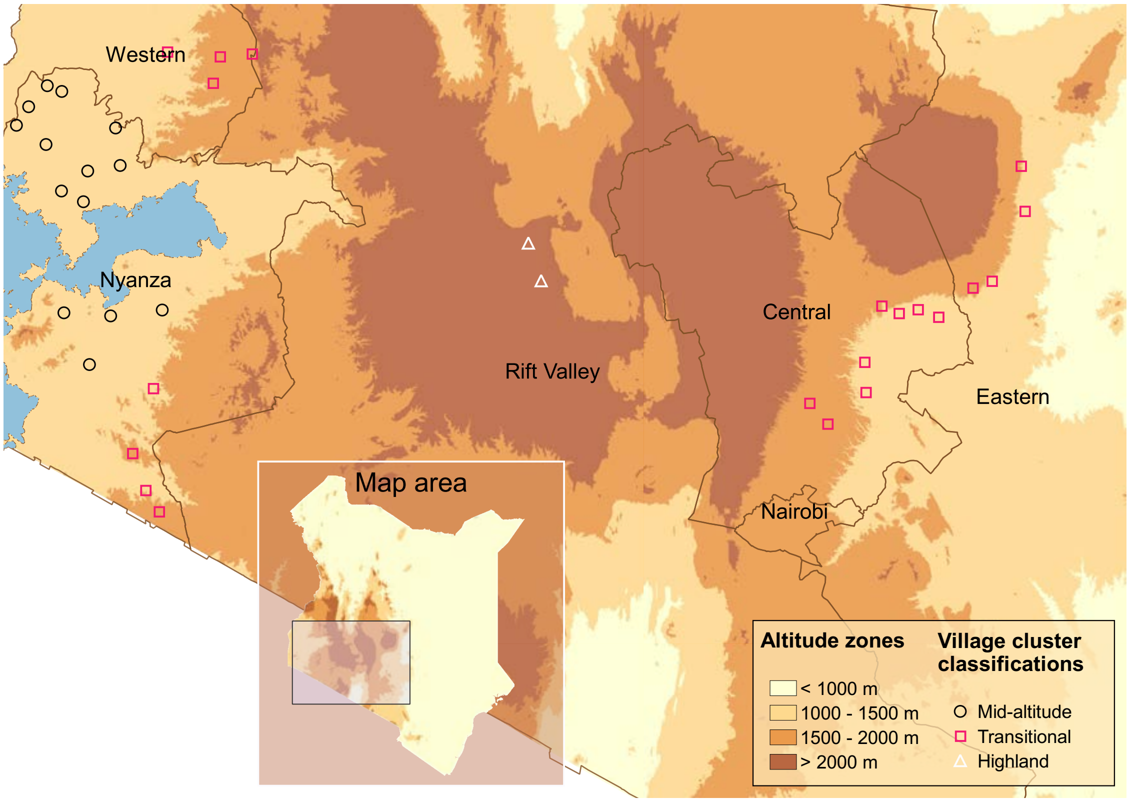


Figure 6: Study sample in western Kenya.

these altitude designations do not completely describe Kenya's agro-ecological zones (see footnote 1), they give insight into the geography of the study area. Each study site is marked by the symbol corresponding to its actual agro-ecological zone classification based on Hassan (1998).

Each site typically contained 3-5 villages within its zone of influence. The easterly, central sites all lie within the transitional zone, while the western sites are divided between the mid-altitude, transitional, and highland zones. For purposes of the analysis that follows, the 2 highland sites were grouped with the western transitional zone sites. The 36 sites randomly selected for the study were grouped into matched pairs based on physical proximity, altitude, and climate. One member of each pair was then allocated to the seed treatment and one to control status. The timeline in Figure 7 displays the full life of the intervention and study. A random sample of 50 farmers was selected for interview in each site, resulting in a total sample size of 1800 farm households.

The seed treatment consisted of three components, with each one following Western Seed's standard seed marketing practice. The first component was the establishment of a nearby demonstration plot for the 2013 main maize season[19] so that farmers could observe the performance of the Western Seed varieties. The second component was the provision of a small sample packet to farmers to try on their own farm for that same season. A Western Seed marketing representative visited each community to distribute the packets and provide further information on the Western Seed hybrids. Sample packets had 250 grams of seed, enough seed to plant one-fortieth of an acre. Farmers were asked to plant their trial packet separately and to keep track of its performance, which most did (see Tjernstrom, 2017). Given the small size of the seed packet, we expected it to inform farmers' future planting decisions, but not to influence their yields or income for the 2013 maize growing season. The third and final component was the offer to pre-order Western Seed hybrids and have them delivered to their village prior to the 2015 maize season. As discussed below, this third element was added following low uptake of Western Seed seed varieties in 2014.

In addition to the core seed treatment, we also implemented a fertilizer intervention for the 2014 maize season that gave fifty kilograms of high-quality fertilizer to randomly selected farmers in both treatment and control sites in the western study areas. At each site, the research team held a public lottery amongst survey participants, with half receiving fertilizer and the others a token gift of cell phone time. The motivation behind this ancillary intervention was to test the claim that the yield gains with Western Seed varieties are much greater for farmers applying fertilizer. We did not implement the fertilizer intervention in the central study area as baseline fertilizer use was quite high.

In summary, in the western study areas, assignment to the treatments randomly divided a total of 1200 farm households into four equally-sized groups:

1. A *Control* group;

---

[19] The primary maize growing season in western Kenya stretches from March to September; some farmers also plant a second maize crop in October, although this second season is typically less productive and receives fewer inputs from most farmers who plant it. The primary maize growing season in central Kenya is the October planting, with a less productive season from March to September.

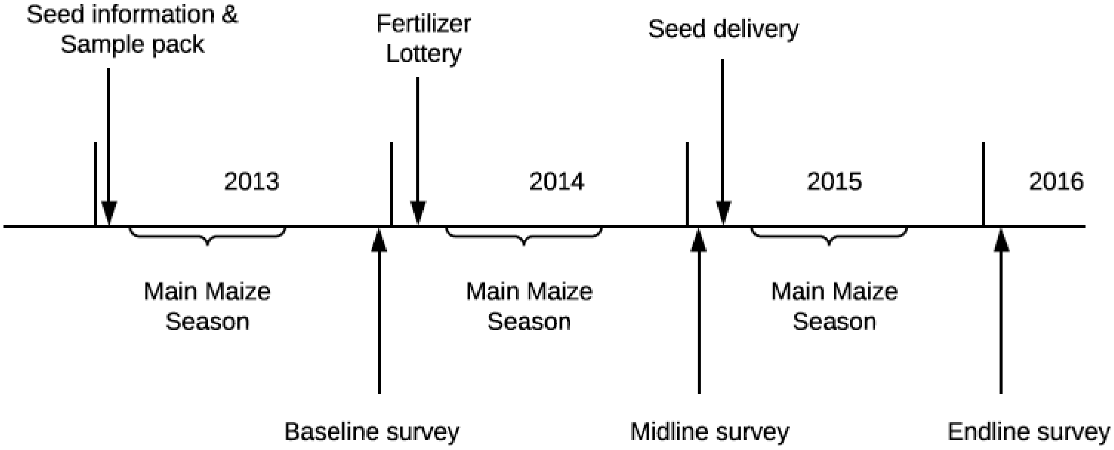


Figure 7: Study timeline.

2. A *Seed Only* treatment group that received Western Seed information and free 250 gram trial seed packets in 2013 and an option to have WSC delivered (at cost) to their home in 2015;

3. A *Fertilizer Only* group that received fertilizer in 2014, but received no seed treatment;

4. A *Seed and Fertilizer* group that received each of the treatments received by groups 2 and 3.

In the central study area, 600 households were divided evenly between groups 1 and 2 only. As shown in Figure 7, baseline, midline and endline surveys were held in both the central and western areas following the 2013, 2014 and 2015 main maize season harvests for western Kenya.

3.3 Baseline Characteristics and Compliance

Table 2 summarizes baseline characteristics of households in each zone: mid-altitude, transitional, and central.[20] We restrict the sample to only those households that reported planting

---

[20] The sample includes 1200 households in western Kenya, of which 700 are in the mid-altitude zone, 400 are in the transitional zone, and 100 are in the highland zone. Given the small sample in the highland zone and the highland zone's similar growing conditions and history of hybrid adoption to the transitional zone, we include observations from the highland zone in our sample from the transitional zone in western Kenya.

maize in each year of the study, as we do in the subsequent empirical analysis. We summarize baseline characteristics in levels for ease of interpretation.

Table 2: Baseline characteristics and compliance.

	Mid-Altitude			**Transitional**	
	All	*Non-Users*	*Users*	*Western*	*Central*
% Main Maize Seasons used Hybrids	26%	11%	96%	84%	83%
% Main Maize Seasons used Fertilizer	26%	20%	56%	81%	92%
Dry maize yield (kg/acre)	234	211	342	553	428
Acres Farmed in Total	1.65	1.58	1.97	1.92	1.29
Acres Planted to Maize	1.32	1.29	1.44	1.32	0.76
Income per-capita (100 Kenya Sh.)	228	210	305	351	507
% Net Income from Ag	34%	37%	22%	36%	69%
% Gross Ag Income from Maize	43%	42%	44%	39%	22%
Average Poverty Probability	32%	33%	26%	32%	13%
% Food insecure	66%	70%	47%	63%	44%
Compliance: % Using Western Seed					
2014 (Midline), Treated	16%	12%	33%	25%	5%
2014 (Midline), Control	1%	1%	2%	16%	1%
2015 (Endline), Treated	20%	18%	33%	31%	10%
2015 (Endline), Control	2%	1%	8%	17%	0%
Observations	589	482	104	428	508

Notes: "Non-Users" are households that did not plant hybrids in at least 4 of the 5 pre-study main maize seasons and "Users" are households that planted hybrids in at least 4 of the 5 pre-study main maize seasons.

The first column summarizes characteristics of households in the mid-altitude zone. Prior to Western Seed's expansion, on average households planted any hybrids in only 26% of previous 5 main maize seasons, and used fertilizer at the same low rate. Unsurprisingly, maize yields also are low, at 234 kilograms per acre. Yet maize is central to the livelihoods of smallholder households in the region; in our sample, households on average plant maize on 80% of their land. Maize is an important source of income and food for smallholder households. In the mid-altitude zone, average annual income per capita is low, at 22,800

Kenyan shillings, or approximately 228 USD, with agriculture contributing 34%, and maize contributing to 43% of agricultural income. This measure of low household income in the mid-altitude zone is substantiated by a separate asset-based index indicates that, on average, a household has almost a 1/3 probability of living on less than 1.25 USD per person per day. Food insecurity is common among households in the sample, with almost 2/3 of households being food insecure at some point during the year.

The theory in Section 2 above suggests that absent well-adapted hybrids, the small farming population will bifurcate into a group using exclusively local seeds and another group relying on hybrids and applying fertilizer. The second and third columns of Table 2 divide up mid-altitude respondents based on their pre-intervention use of hybrids. Hybrid users are those who had planted hybrids in at least 4 of the 5 pre-study main maize seasons. Hybrid non-users are those who planted hybrids in less than 4 of those seasons.[21] As can be seen in the table, the smaller hybrid-users group almost exclusively relies on hybrids, applying fertilizer almost 60% of the time. The larger non-users group rarely uses hybrids and uses fertilizer at about one third rate of the users group. As expected, yields are much higher at baseline for the hybrid users group, at 342 kilograms per acre compared with the non-user group average of 211 kilograms per acre.[22] Incomes for the users group is about 50% higher, due in large part to greater non-agricultural income, although food security and predicted poverty rates are relatively similar across the users and non-users groups.

Columns 4 and 5 of Table 2 report characteristics of farms in the higher altitude transitional zone in both western and central Kenya. Consistent with Figure 1, hybrid and fertilizer use are uniformly high and maize yields top those of even hybrid users in the mid-altitude zones, as would be expected given that existing hybrids are better adapted to this agro-ecological zone. Incomes are higher and poverty indicators are lower than in the

---

[21] Of the 700 households in the mid-altitude zone, we categorize 125 as users of hybrids in the seasons preceding the study. This includes four households that planted maize in fewer than 4 of the 5 pre-study main maize seasons, but planted hybrids in each of the seasons in which they planted maize.

[22] Two mid-altitude households have missing data for past hybrid use, which is why the observations for the non-user and user samples do not sum to the total number of observations in the mid-altitude sample.

mid-altitude zone, especially for farmers in the central region who are able to grow coffee and other cash crops and who correspondingly get much less of their agricultural income from maize. While farmers in the transitional zone are largely better-off than those in the mid-altitude zones, material well-being is still low relative to global standards.

The lower portion of Table 2 reports compliance with the seed treatment, defined as the percentage of farmers who responded to the informational treatment and purchased a Western Seed variety during the midline and endline seasons. As can be seen, compliance was modest in 2014. Field reports indicated a number of factors that conspired to to lower uptake that year.[23] These challenges motivated a seed delivery program in 2015. Adoption in 2015 was 5-6 percentage points greater than adoption in 2014 for each of three agro-ecological zones. The different intensity of treatment in 2014 and 2015 of course could lead different types of households to adopt Western Seed hybrids in those years; because of this, in Section 5 we report separate estimates for the endline year.

Finally, our ability to make inferences from the RCT critically depends on the assumption that random assignment of households to treatment groups being uncorrelated household characteristics, both observable and unobservable. To shed light on the validity of that assumption, Appendix C presents balance tables for observable characteristics for our sample. In general, baseline differences between treatment groups are not large in magnitude relative to average baseline levels in the control group. Balance on observables gives us confidence that omitting these variables from our estimation will not bias our treatment effect estimates. However, our treatment effect estimates may be biased due if a baseline variable with modest imbalance across treatment groups is strongly correlated with the dependent variable. Since this is most likely to be true for baseline levels of the dependent variable, we estimate ANCOVA specifications that control for baseline levels of dependent variables. Finally, balance on baseline measures of household characteristics gives us confidence that we also

---

[23] Adoption in 2014 was lower than anticipated due to a number of factors, including the former parastatal Kenya Seed subsidizing its seeds, offering added incentives to agro-dealers to sell Kenya Seed. In addition, Western Seed faced challenges in expanding their seed promotion to new regions like central Kenya.

have balance on household characteristics that we cannot observe.

4 Average Effects on Yields by Agro-Ecological Zone

This section uses data from our RCT to test whether Western Seed's adapted improved seed varieties increase farmer yields. To account for differences in socioeconomic status and maize seed markets across agro-ecological zones, we estimate the effects of Western Seed hybrids separately for each zone. We find large impacts in the mid-altitude zone, but little to no impact in the transitional zones of western and central Kenya.

To identify the impact of Western Seed maize hybrids on maize yields of smallholder farmers in our study areas, we estimate average treatment effects of the random variation in access to Western Seed varieties. In the mid-altitude and transitional zones, following McKenzie (2012) we estimate average treatment effects on maize yields, y_{ivspt}, using a pooled analysis of covariance intention to treat specification.[24] The specification for farm i in village v, site s, matched pair p and time period t is:

$$(2) \qquad y_{ivspt} = \beta_0 + \beta_1 y^0_{ivsp} + \delta_1 I^W_{sp} + \delta_2 I^F_{ivsp} + \delta_3 \left[I^W_{sp} \times I^F_{ivsp} \right] + \left[\mu_p + \varepsilon_{ivspt} \right], t = 1, 2$$

where y^0_{ivsp} is baseline maize yields, I^W_{sp} and I^F_{ivsp} are binary indicators for assignment to the Western Seed and fertilizer treatments respectively, μ_p is a matched pair fixed effect included to account for stratification of the seed treatment following Bruhn & McKenzie (2009). The error term (ε_{ivspt}) is clustered by village, the level of stratification for the fertilizer treatment. In the central region, where there was no fertilizer lottery, we estimate equation 2 after eliminating the terms involving I^F_{ivsp}.

Table 3 reports the estimated coefficients for equation 2 by agro-ecological zone. To lessen

---

[24] Our preferred specification estimates a single set of treatment effects by pooling observations from both of the post-treatment years of observations. When separating observations by these two post-treatment years, treatment effect estimates are qualitatively similar to the estimates from our preferred specification, both for the analysis in this section as well as the heterogeneity analysis in the subsequent section.

the role of outliers, we transform yields using the inverse hyperbolic sine function. The lower half of the table reports the implied percentage effects of the different treatments following the method of Bellemare & Wichman (2019).[25] As reported in the table, for the mid-altitude zone, we estimate that the seed treatment increased average yields by 25%. Given that this is an intention to treat estimate (with a compliance rate of about 20% as reported in Table 2), the actual yield increases experienced by those who adopted Western Seed because of the seed treatment (impact of the treatment on the treated) is approximately five-times greater, indicating substantial yield gains among adopters of Western Seed. Given that the mid-altitude area is one with relatively low hybrid use, impacts of Western Seed likely come from two sources. The first are first generation adoption effects from the intervention inducing households to switch from local varieties to hybrid varieties. The second are second generation adoption effects from the intervention inducing households to switch from other hybrid varieties to Western Seed locally adapted hybrid varieties. In the next section, we explore this issue further to distinguish the magnitudes of the two effects.

In addition to the primary seed treatment effect on yields in the mid-altitude zone, the other striking result in Table 3 is the negative interaction effect between the seed and fertilizer treatments. As discussed in Sub-Section 3.2 above, in-kind grants of fifty kilograms of fertilizer were made to randomly selected study households in order to see if relaxing constraints to fertilizer acquisition might boost returns to the seed treatment. If access to fertilizer had proven not to be a constraint, we might have expected to see the impact of the fertilizer grant, both alone and in combination with the seed treatment, to be zero. Instead, Table 3 shows a positive effect of the direct impact and a negative impact of the interaction effect. In order to better understand this unexpected outcome, we look more closely at the impact of the different treatments on fertilizer use in Appendix D. We find substantially greater midline leakage from fertilizer lottery winners in seed treatment communities than in

---

[25] For example for the seed treatment, the estimated percentage change is calculated as $\exp(\hat{\delta}_1 - 0.5\hat{\sigma}^2_{\hat{\delta}_1}) - 1$, where $\hat{\sigma}^2_{\hat{\delta}_1}$ is the estimated variance of $\hat{\delta}_1$.

Table 3: Effects on maize yield (IHST of kg/ac).

	Mid-Altitude	**Transitional**	
	All Farms	*Western*	*Central*
Seed Treatment, $\hat{\delta}_1$	0.23*	-0.17	0.10
	(0.13)	(0.12)	(0.15)
Fertilizer Treatment, $\hat{\delta}_2$	0.23**	0.11	–
	(0.11)	(0.09)	–
Seed*Fertilizer Treatment, $\hat{\delta}_3$	-0.23*	0.20	–
	(0.14)	(0.12)	–
Baseline Yield, $\hat{\beta}_1$	0.11***	0.15***	0.20***
	(0.03)	(0.02)	(0.05)
Percent effects			
Seed Treatment	25%	-17%	9%
Fertilizer Treatment	26%	11%	–
Control mean	5.97	6.86	6.27
Observations	1178	856	1016
R-squared	0.07	0.06	0.11

Pair indicator variables included as controls.
Standard errors clustered by village.
* = 10% significance, ** = 5%, *** = 1%

non-seed treatment communities, consistent with the interpretation that there was greater rates of social taxation of “windfall fertilizer” in treatment sites. Given this configuration, it is unsurprising that households that benefited from both the seed treatment and the fertilizer lottery basically only benefited from the seed treatment, with the added benefit of fertilizer largely erased. This pattern of differential social taxation across treatment and control villages is thus consistent with the pattern of estimated treatment effects shown in Table 3.

While we do see significant impacts of the seed treatment in the mid-altitude zone, in the transitional zone the estimated treatment effects on yields are relatively small and do not differ from zero at the five percent significance level. Recall that hybrid use was already quite high in this zone prior to the study, and that net compliance was also quite low (see Table 2).

5 Heterogeneous Effects on Yields by Past Hybrid Use

As reported above, we find substantial yield impacts of Western Seed varieties in the mid-altitude zone. Given that prior adoption of hybrids in this area was modest, the question remains whether the observed impacts reflect the fact that Western Seed varieties outperform local varieties as well as other commercially available hybrids (largely bred for higher altitude maize-growing areas). The simple theoretical model developed in Section 2 is silent on the question as to which group will benefit most from a well-adapted improved seed: those who make the switch from local, unimproved seed or those that switch from less well-adapted improved seed. To gain empirical purchase on this question, this section splits our mid-altitude sample between households that consistently used hybrids prior to the study period and those that did not (see Table 2).

Table 4: Heterogeneous impact of locally adapted improved seeds in the mid-altitude zone (IHST of kg/ac).

	Mid-Altitude Zone			
	Hybrid Non-Users		*Hybrid Users*	
	Pooled	Endline	Pooled	Endline
Seed Treatment, $\hat{\delta}_1$	0.20	0.26*	0.40**	0.78***
	(0.15)	(0.15)	(0.16)	(0.21)
Fertilizer Treatment, $\hat{\delta}_2$	0.22	0.18	0.35**	0.54**
	(0.13)	(0.14)	(0.16)	(0.21)
Seed*Fertilizer Treatment, $\hat{\delta}_3$	-0.21	-0.21	-0.25	-0.44*
	(0.17)	(0.18)	(0.24)	(0.26)
Baseline Yield, $\hat{\beta}_1$	0.12***	0.09**	0.02	-0.02
	(0.03)	(0.04)	(0.05)	(0.05)
Percent effect				
- Seed Treatment	21%	29%	47%	113%
- Fertilizer Treatment	23%	18%	39%	68%
Control mean	5.92	6.03	6.13	5.98
Observations	964	482	208	104
R-squared	0.06	0.03	0.19	0.22

Pair indicator variables included as controls.
Standard errors clustered by village.
Users planted hybrids in 80-100% of main seasons (2007-2012).
* = 10% significance, ** = 5%, *** = 1%

Table 4 displays the ITT estimates for equation 2 for the hybrid users and non-users sub-samples. As shown in the bottom panel of the table, the estimated impacts of the seed treatment are 21% for the non-user group and 47% for the user group, with only the latter estimate attaining conventional levels of statistical significance. As shown in Table 2, net compliance for the users group was 31% in the midline and fell to 25% at the time of the endline (with the decline driven by a substantial uptick in WSC seed use by the control group). For the non-users group, net compliance was only 11% in midline, rising to 17% in the endline following the ancillary seed delivery intervention. This lower level of compliance explains at least in part the smaller magnitude and statistical insignificance of the ITT estimate for the non-users group. For the users group, the estimated 47% ITT impact translates into a robust impact for those that actually adopted the seeds.

Given the compliance issues that especially surrounded the midline year, the research team implemented an ancillary seed delivery service, as described in Section 3.3 above. While the delivery was less effective than anticipated, it did boost compliance among the non-users group.[26] The second and fourth columns of Table 4 thus present separate estimates for the endline year data only. As can be seen, for the non-users group, the estimated impact rises and becomes statistically significant and indicates an ITT treatment effect of a 29% yield increase. The estimated impact for the hybrid users group retains its significance and jumps substantially to an implied 113% yield increase. The fact that the impacts on the level of yields for the baseline hybrid user population are larger than for the non-user population (even after accounting for differences in the net compliance rate) would seem puzzling at first glance. However, this finding is consistent with an interpretation that the ability of the more resource-constrained non-user farm households to garner the full genetic benefits of Western Seed hybrids bred for their agro-ecological zone.

Because we derive our results from theory for farm households whose key outcome is net

---

[26] Phone orders for the seed were quite high, with 55% of treatment farmers ordering seed. Unfortunately, at time of delivery, only 16% of treatment farmers bought seed via the delivery intervention. Many farmers proved unwilling or unable to pay for the seed despite the fact that it had been made clear that the seeds were being sold at the market price and not given away for free.

income, we use our estimates of impacts on yields to calculate the change in net revenues that two stylized types of mid-altitude farmers (hybrid non-users and hybrid users) would receive from adopting locally adapted Western Seed hybrids. Adjusting the Table 4 ITT estimates for the compliance rates, we calculate that "non-user" farmers who switched from local seeds would have experienced a 339 kilogram increase in production per-acre (a 161% increase), while farmers who switched from other hybrids and used fertilizers would have experienced a 482 kilogram increase per-acre (a 141% increase).[27] Using market prices for outputs, seeds and fertilizer, these increases imply a 130% increase in net revenues for non-user farmers switching from local seeds, and a 245% increase for user farmers switching from NLA hybrids to WSC varieties.[28] These financial figures place the impacts squarely in Case 1 of Table 1 above, with both an intercept and a slope effect and both low resource and better resourced farmers benefit from the Western Seed adaptive breeding program.

6 Conclusion

While the average Kenyan farmer deviates from the sub-Saharan African pattern of low hybrid maize adoption, the mid-altitude zone of Kenya more closely resembles the rest of the continent: persistently modest hybrid use over the last 50 years. A simple theory of the supply and demand for hybrid seeds suggests a rationale for this persistent pattern. Large-scale seed companies do not supply well-adapted hybrid varieties to small markets. Neither do local seed companies who may enjoy informational (and, perhaps, cost-) advantages, but are constrained by their lack of capital. Given the lack of supply of adapted varieties, our model suggests that adoption of the available (non-adapted) hybrids will be low and restricted to better-resourced farmers within the niche market. Absent locally-adapted hybrids, farmer productivity and incomes stagnate, especially among the poorest farmers.

---

[27] For these calculations, we use the pooled ITT estimates from Table 4 for both the hybrid users and non-user groups and the compliance rates reported in Table 2. These treatment on the treatment percent change in yield are then applied to the baseline yield levels, also reported in Table 2.

[28] Consistent with the survey data, we assume that the non-user group applies on inorganic fertilizer, while the user group applies 50 kilograms per acre, the median value for this group in the sample.

An external infusion of impact investor capital to a local seed company based in Kenya's mid-altitude zone allowed us to explore the impacts of disrupting this seed market equilibrium. By 2010, Western Seed had developed and registered hybrid maize varieties—building on publicly-provided parent seed lines from CIMMYT—and were supplying high-performing locally-adapted seed varieties to the mid-altitude niche market. The impact investment capital allowed the company to rapidly expand its seed production capacity and expand its marketing to the mid-altitude and the (higher) transitional zone. In collaboration with Western Seed, we established a three-year RCT to study the impact of offering the new seed varieties.

The key findings from the empirical analysis are that the new varieties offered substantial benefits to farmers located in the mid-altitude region. On average, the ITT estimate of the yield impact implies a 25% yield increase. Given the modest compliance rates, the actual impact on adopters is estimated to be five times that amount. In the higher-altitude zones where Western Seed also expanded, the yield benefits are both economically and statistically negligible, consistent with the expectation that the seed sector has already adapted varieties for these larger and more lucrative zones.

Consistent with our theoretical model, the data reveal that prior to the experimental introduction of Western Seed hybrid, farmers in the mid-altitude zone could be divided into two groups: a larger group (82% of the sampled farmers) who almost never use improved seeds and rarely use fertilizer, and a smaller group (18%) who almost always use improved seeds and fertilizer. While our statistical power suffers as we sub-divide the mid-altitude farmers into these two groups, we estimate that the yield impacts on farmers already using hybrids are actually larger than those on the farmers who switched from local, unimproved seed varieties. In the last year of the RCT, we estimate that the productivity increases for those who actually adopted Western Seed are on the order of 100-150% for those who switched from local seeds, and 150% for prior hybrid adopters who are able to apply fertilizers to their maize. While large, these figures are in line with the seed company's claims about the yield potential of their varieties in the mid-altitude region.

Stepping back, these results indicate that the seed system was leaving substantial productivity-enhancing opportunities on the table. While our data do not allow us to test the performance of different seed systems *per se*, the patterns that we observe are consistent with a theoretical model in which small agro-ecological niches will remain underserved absent the hybrid partnerships (public investments in foundation seed and impact investment capital) that allowed Western Seed to expand. Our results cautiously suggest that such a hybrid system has much to offer other areas of sub-Saharan Africa where cereal yields lag even further behind.[29] These observations do not say that the other constraints discussed in the introduction do not matter. However, better adapted and more profitable improved varieties would alter incentives for risk averse or even time-inconsistent farmers to adopt them.

Entry of new actors into niche breeding may transform innovation markets for agricultural technologies in regions like mid-altitude Kenya to more closely resemble the present in regions like transitional and highland Kenya, where many firms compete in hybrid development. Such a change would shift research priorities toward studying the implications of competition and product differentiation between firms on the productivity and welfare of agricultural households in these environments. From this perspective, our finding that locally well-adapted varieties offer important benefits to both poor and better-resourced farmers indicates that such adaptive breeding, whatever its source, can offer substantial economic and social benefits.

---

[29] However, we should stress that our modeling assumes that local companies have lower fixed costs of hybrid development due to greater knowledge of the local growing conditions. While we believe this assumption approximates the case of hybrid development in Kenya in the early 2000s, recent advances in big data analysis may soon tilt cost advantages in favor of large multinational firms that can realize economies of scope from drawing on genetic markers and data from agronomic test trials in agro-ecological environments around the world. While development of hybrids by Western Seed arose from public investments by CIMMYT's breeding program that drew on its international scope of research, in the future multinational companies may realize even greater economies of scope independently, raising the potential of these large firms entering the market for niche breeding at low cost.

Acknowledgements

We are grateful for comments from Jim Gaffney, GianCarlo Moschini, Wendong Zhang, and seminar participants at Iowa State University, the University of California–Davis, and the University of Wisconsin–Madison. We thank our survey respondents for their enthusiasm, generosity, and patience during the course of this research.

Funding: This work was supported by Acumen, a non-profit impact investment firm; the American people through the United States Agency for International Development Cooperative Agreement No. AID-OAA-L-12-00001 with the BASIS Feed the Future Innovation Lab; and the Agricultural Technology Adoption Initiative (ATAI) administered by JPAL at MIT and the Bill and Melinda Gates Foundation. The contents are the responsibility of the authors and do not necessarily reflect the views of Acumen, USAID, the US Government, or other funders. Declarations of interest: none.

References

Abay, Kibrom, Blalock, Garrick, & Berhane, Guush. 2017. Locus of control and technology adoption indeveloping country agriculture: Evidence from Ethiopia. *Journal of Economic Behavior and Organization*, **143**(November), 98–115.

Bellemare, Marc F., & Wichman, Casey J. 2019. Elasticities and the Inverse Hyperbolic Sine Transformation. *Oxford Bulletin of Economics and Statistics*, **82**(1), 50–61.

Bruhn, Miriam, & McKenzie, David. 2009. In Pursuit of Balance: Randomization in Practice in Development Field Experiments. *American Economic Journal: Applied Economics*, **1**(4), 200–232.

Byerlee, Derek, & Traxler, Greg. 2001. The Role of Technology Spillovers and Economies of Size in the Efficient Design of Agricultural Research Systems. *Chap. 9, pages 161–186 of:* Alston, Julian M., Pardey, Philip G., & Taylor, Michael J. (eds), *Agricultural Science Policy: Changing Global Agendas*. Baltimore and London: The Johns Hopkins University Press.

Carter, Michael R., Laajaj, Rachid, & Yang, Dean. forthcoming. Temporary Subsidies and Green Revolution Technology Adoption by Farmers and Their Social Networks. *American Economic Journal: Applied Economics.*

Dercon, Stefan, & Christiaensen, Luc. 2011. Consumption risk, technology adoption and poverty traps: evidence from Ethiopia. *Journal of Development Economics*, **96**(2), 159–173. working paper.

Duflo, Esther, Kremer, Michael, & Robinson, Johnathan. 2011. Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya. *American Economic Review*, **101**(6), 2350–2390.

Evenson, R. E., & Gollin, D. 2003. Assessing the impact of the Green Revolution, 1960 to 2000. *Science*, **300**(5620), 758–762.

Fund, Acumen. 2010. *Acument Fund announces investment in Western Seed Company Limited.* Tech. rept.

Gerhart, John. 1975. *The diffusion of hybrid maize in Western Kenya—Abridged by CIMMYT.* Tech. rept. Centro Internacional de Mejoramiento de Maiz y Trigo, Mexico City.

Griliches, Zvi. 1957. Hybrid Corn: An Exploration in the Economics of Technological Change. *Econometrica*, **25**(4), 501–522.

Griliches, Zvi. 1960. Hybrid Corn and the Economics of Innovation. *Science*, **132**(3422), 275–80.

Hassan, R M. 1998. *Maize Technology Development and Transfer: A GIS Application for Research Planning In Kenya.* Wallingford and New York: CAB International.

Kantor, Shawn, & Whalley, Alexander. 2019. Research Proximity and Productivity: Long-Term Evidence from Agriculture. *Journal of Political Economy*, **127**(2).

Karlan, Dean, Osei, Robert, Osei-Akoto, Isaac, & Udry, Christopher. 2014. Agricultural Decisions after Relaxing Credit and Risk Constraints. *Quarterly Journal of Economics*, 597–652.

KEPHIS. 2018. *National Crop Variety List - Kenya*. Tech. rept. Kenya Plant Health Inspectorate Service, Nairobi.

Laajaj, Rachid, & Macours, Karen. forthcoming. Measuring Skills in Developing Countries. *Journal of Human Resources*.

Langyintuo, Augustine S., Mwangi, Wilfred, Diallo, Alpha O., MacRobert, John, Dixon, John, & Bänziger, Marianne. 2010. Challenges of the maize seed industry in eastern and southern Africa: A compelling case for private-public intervention to promote growth. *Food Policy*, **35**(4), 323–331.

Marenya, P., & Barrett, C. B. 2009. State-conditional Fertilizer yield response on western Kenyan farms. *Amer. J. Agr. Econ.*, **91**(4), 991–1006.

McKenzie, David. 2012. Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics*, **99**(2), 210–221.

Partners, Pearl Capital. no date. *Impact investment Understanding financial and social impact of investments in East African agricultural businesses*. Tech. rept. The Gatsby Charitable Foundation.

Spielman, David J, Kolady, Deepthi, Cavalieri, Anthony, & Rao, N Chandrasekhara. 2014. The seed and agricultural biotechnology industries in India: An analysis of industry structure, competition, and policy options. *Food Policy*, **45**, 88–100.

Suri, Tavneet. 2011. Selection and Comparative Advantage in Technology Adoption. *Econometrica*, **79**(1), 159–209.

TAMPA2. 2004. *Tegemeo Agricultural Monitoring and Policy Analysis Project.* Tech. rept. Egerton University - Tegemeo Institute of Agricultural Policy and Development and Michigan State University, Nairobi.

TAPRA. 2010. *Tegemeo Agricultural Policy Research Analysis Project.* Tech. rept. Egerton University - Tegemeo Institute of Agricultural Policy and Development and Michigan State University, Nairobi.

Tjernstrom, Emilia. 2017. *Learning from Others in Heterogeneous Environments.*

Appendix A: Linear Returns to Fertilizer

As Section 2 describes, standard agronomic practice recommends that farmers who cannot apply the optimal amount of fertilizer to their entire crop, should apply the optimum rate to a portion of their crop rather than applying fertilizer at a diluted rate to the entire crop. Using a simple economic model that has roots in the efficiency wage literature's portrayal of the response of human work capacity to nutrition (see **?**), this appendix shows that optimized returns to fertilizer will be linear under profit maximization whenever the farmer is constrained in her purchase of fertilizer.

Figure 8 graphs the additional production per-hectare (q) that occurs as fertilizer intensity (f) increases under the assumption that marginal returns to fertilizer exhibit increasing returns over a range and then diminishing returns thereafter. Without fertilizer, farmers receive a fixed amount α_0, and $q(f)$ is the additional output received in addition to the base, no fertilizer amount. In the figure, the slope of the ray from origin is the average product of fertilizer (q/f), or the bang for the buck spent on fertilizer. As can be seen by visual inspection, f^* is the fertilizer intensity that maximizes additional production per-unit fertilizer given the S-shaped production function. No other intensity will give more. We denote the slope of the ray that intersects the function q at input level f^* as α_1.

Consider the case where $f^* = 100$ kg/ha and a farmer with one hectare of maize has only 50 kg of fertilizer, then she will maximize output and income by concentrating the 50 kg on 0.5 hectare and putting zero fertilizer on the rest. As can easily confirmed visually, alternative allocations (*e.g.*, 50 kg/hectare on her entire maize plot, yielding a fertilizer intensity of $\frac{f^*}{2}$) will yield less output for the same input expenditure because it fails to fully exploit the increasing returns portion of the returns to fertilizer function, $q(f)$.[30] That is,

---

[30] This argument is exactly identical to the initial contributions to the efficiency wage literature in which an employer (who is indirectly buying nutrition by paying workers a wage) will never pay a worker less than the efficiency wage because worker productivity falls off more quickly than cost when descending the increasing returns portion of the efficiency wage curve.

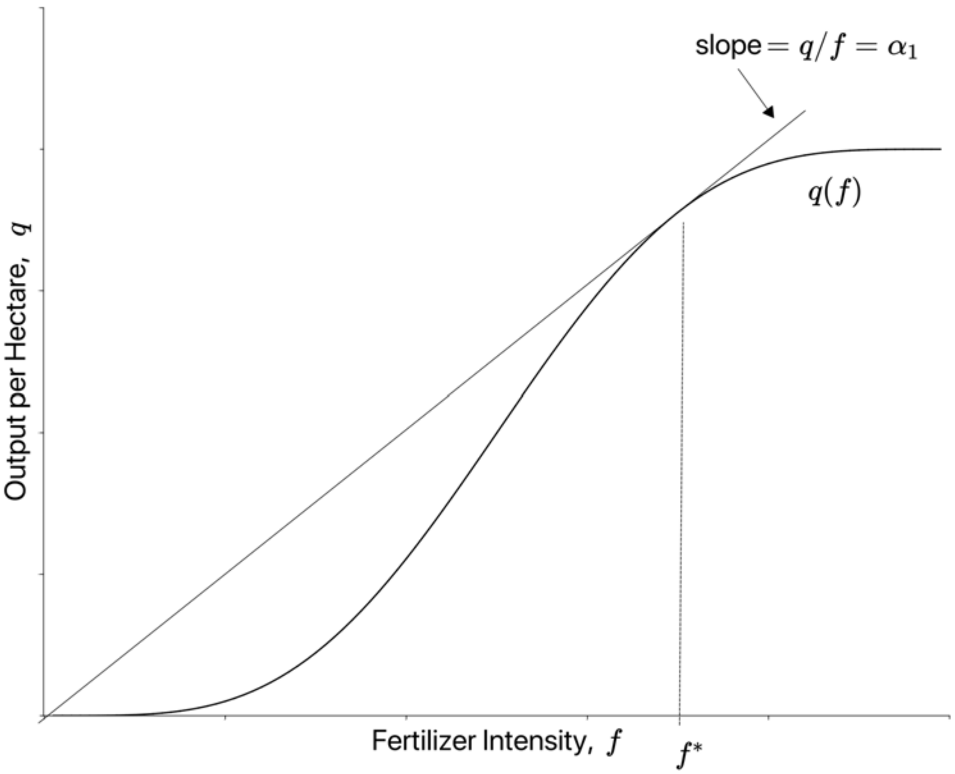


Figure 8: Returns to Fertilizer

for this case of constrained access to fertilizer:

$$0.5q(0) + 0.5q(f^*) > 0.5q(\varepsilon) + 0.5q(f^* - \varepsilon), \forall \varepsilon.$$

For this case, when fertilizer is optimally applied, marginal returns to additional units of fertilizer will always be a constant α_1, not because marginal returns do not change, but because the farmer optimally adjusts fertilizer application to obtain the constant, maximal return. Once the farmer can apply fertilizer at intensity f^* on all her land, further use of fertilizer will face diminishing returns as beyond that level, the function $q(f)$ is strictly concave.

This intuition can be captured by the following profit maximization problem:

$$(3) \quad \underset{F,T_f}{Max} \quad Q - p_f\bar{F}$$

$$subject\ to:$$

$$Q = T_f q(F/T_f) + \left(\bar{T} - T_f\right) q(0)$$

$$(4) \quad p_f F \leq K$$

$$T_f < \bar{T}$$

where p_f is the price of fertilizer and K is the working capital available to purchase fertilizer. Note that this simple specification allows any number of possible outcomes, including zero use of fertilizer ($F, T_f = 0$), which will occur when $\alpha_1 < p_f$; and, use of fertilizer in excess of f^* ($T_f = \bar{T}$ and $F > f^*\bar{T}$), which will happen when $K > p_f\bar{T}f^*$ and $\alpha_1 > p_f$. As discussed in Sub-Section 2.1, the relevant case for low income small scale farmers, who cannot borrow against their future income to buy fertilizer, is that $K < p_f\bar{T}f^*$. For these farmers, the *optimized* production function they face can thus be written as $\alpha_0 + \alpha_1 f$, with constant marginal returns to fertilizer. As further analyzed in Sub-Section 2.1, variety choice and the decision whether or not to use any fertilizer will be based on those linear returns for liquidity constrained farmers.

Appendix B: Proofs of Results on the Demand for Improved Seed Varieties

First we demonstrate that, for sufficiently poor households, optimized utility conditional on minimal adoption of the NLA improved variety will be strictly less than the non-adoption alternative ($\triangle^{nr} < 0$ for small z_1). For an arbitrary and small value ϵ, we can construct a level of cash-on-hand:

$$\underline{z_1} \equiv \underline{H} s^n + (\bar{H} - \underline{H}) s^r + \epsilon$$

The corresponding utility conditional on minimal adoption of the NLA improved variety is:

$$V^n(\underline{z_1}, \bar{H} | H^n = \underline{H}) = u(\epsilon) + \beta u(\bar{H}\alpha_0^n - \underline{H}(\alpha_0^r - \alpha_0^n))$$

where we have assumed that the return from planting the retained variety without fertilizer is strictly greater than the return from planting the NLA variety without fertilizer ($\alpha_0^r - \alpha_0^n > 0$).

The corresponding utility under the non-adoption alternative is:

$$V^l(\underline{z_1}, \bar{H} | H^n = H^a = 0) = u(\underline{H}(s^n - s^r) + \epsilon) + \beta u(\bar{H}\alpha_0^n)$$

which is strictly greater than utility conditional on minimal adoption of the NLA improved variety, since we have assumed that NLA seed varieties are more expensive than retained seed varieties ($s^n > s^r$).

Second, we demonstrate that the difference in the value functions increases with initial cash on hand. The difference in value functions varies with initial cash on hand as follows:

$$\frac{\partial \Delta^{nr}}{\partial z_1} = u'(.)\underline{H}\frac{\partial f^n}{\partial z_1}(\beta \alpha_1^n - p^f)$$

Thus the difference in the value functions increases with initial cash on hand when two conditions are met. First, returns to fertilizer use on the NLA variety must be large relative to the discount rate ($\frac{\alpha_1^n}{p^f} > \beta$), which we assume to true. Second, fertilizer use on the NLA variety must increase with initial cash on hand, which we formally derive by applying the implicit function theorem to the household's first-order condition:

$$\frac{df^n}{dz_1} = \frac{u''(.)\underline{H}p^f}{u''(.)((\underline{H}p^f)^2 + \beta(\underline{H}\alpha_1^n)^2)} > 0$$

Finally, given that the difference in value functions and the level of fertilizer use increase with initial cash on hand, we can define a critical cash on hand level and corresponding level of fertilizer use ($\tilde{z}^n, \tilde{f}^n$) at which adoption of the NLA improved variety begins. Households with cash on hand greater than this critical level will adopt the NLA improved variety and apply more fertilizer than the critical level. Households with cash on hand less than this critical level will not adopt the NLA improved variety.

Table 5: Balance at baseline, Western (N=1017)

	Summary Stats		Estimates from OLS		
	(1)	(2)	(3)	(4)	(5)
	Pooled	Control	Seed	Fert	Seed*Fert
Hybrid main seasons (0-1)	0.51	0.52	0.00	-0.06**	0.06
	(0.44)	(0.45)	(0.04)	(0.03)	(0.04)
Fertilizer main seasons (0-1)	0.49	0.53	-0.02	-0.07**	0.03
	(0.46)	(0.46)	(0.04)	(0.03)	(0.05)
Dry maize yield (kg/ac)	368.49	375.01	7.50	-12.75	-18.17
	(425.64)	(428.83)	(37.21)	(26.93)	(40.24)
Acres (maize)	1.32	1.38	-0.26***	0.03	0.19
	(1.17)	(1.32)	(0.09)	(0.10)	(0.12)
Acres (total)	1.76	1.82	-0.27**	0.06	0.16
	(1.51)	(1.59)	(0.12)	(0.12)	(0.15)
Income per capita (100 ksh)	279.60	252.91	25.31	10.52	38.22
	(391.37)	(347.63)	(34.66)	(30.42)	(43.33)
Poverty (0-1)	0.32	0.33	-0.02	-0.00	0.01
	(0.23)	(0.24)	(0.02)	(0.02)	(0.03)
Dietary diversity (0-12)	6.68	6.52	0.28**	0.07	-0.03
	(1.62)	(1.65)	(0.14)	(0.14)	(0.19)
Food insecure (0/1)	0.65	0.70	-0.07**	-0.03	0.00
	(0.48)	(0.46)	(0.04)	(0.04)	(0.06)

Pooled and Control report means (standard deviations). Seed, Fert, and Seed*Fert report point estimates obtained by OLS with pair indicators as controls (standard errors clustered by village). Significance: * = 10%, ** = 5%, *** = 1%

Appendix C: Balance Checks in Western and Central

Table 5 estimates how baseline characteristics differ by treatment status in Western. Table 6 estimates how baseline characteristics differ by treatment status in Central.

Table 6: Balance at baseline, Central (N=508).

	Summary Stats		OLS
	(1)	(2)	(3)
	Pooled	Control	Seed
Hybrid seasons (0-10)	7.78	7.82	-0.06
	(3.55)	(3.49)	(0.36)
Fertilizer seasons (0-10)	8.65	8.64	0.02
	(2.91)	(2.94)	(0.35)
Dry maize yield (kg/ac)	428.46	432.64	-5.14
	(489.72)	(515.44)	(39.63)
Acres (maize)	0.76	0.74	0.04
	(0.70)	(0.70)	(0.08)
Acres (total)	1.29	1.27	0.04
	(1.14)	(1.08)	(0.14)
Income per capita (100 ksh)	507.28	424.16	167.74**
	(738.67)	(552.79)	(70.15)
Poverty (0-1)	0.13	0.13	-0.01
	(0.17)	(0.18)	(0.01)
Dietary diversity (0-12)	7.63	7.63	-0.01
	(1.45)	(1.40)	(0.15)
Food insecure (0/1)	0.44	0.46	-0.03
	(0.50)	(0.50)	(0.05)

Pooled and Control report means (standard deviations). Seed reports point estimates obtained by OLS with pair indicators as controls (standard errors clustered by village).

Significance: * = 10%, ** = 5%, *** = 1%

Appendix D: The Fertilizer-Seed Interaction Puzzle

In order to explore to the impact of fertilizer treatment, we define a set of four mutually exclusive dummy variables: D^S indicates those farms that were assigned to the seed treatment but lost the fertilizer lottery; D^F are farms that won the fertilizer lottery but were not assigned to the seed treatment; and, D^{SF} are farms that were assigned to the seed treatment and won the fertilizer lottery. Using these terms, we can write the following regression for the amount of fertilizer used by farm household i in matched pair p in time period t, where $t = 1, 2, 3$ covers the baseline, midline and endline survey rounds:

$$f_{ist} = \sum_{t=1}^{3} \left(\alpha_t^C + \alpha_t^S D_{ist}^S + \alpha_t^F D_{ist}^F + \alpha_t^{SF} D_{ist}^{SF} \right) + \mu_p + \varepsilon_{ist}$$

Under this specification, the α_t^C parameters capture the fertilizer use by the pure control group (farm households that were not assigned to the seed treatment and did not win the fertilizer lottery). The other parameters capture the additional effect of being in one of the three mutually exclusive treated groups at baseline, midline and endline. Estimation included a matched pair fixed effect and standard errors were clustered at the village level.

Table 7 gives the parameter estimates for this model. The estimates evidence substantially greater midline leakage from fertilizer lottery winners in seed treatment communities than in non-seed treatment communities. Looking at total use of fertilizer on all crops on the farm, we find that in control sites without the seed treatment, the average winner of the fertilizer lottery increased use of fertilizer in 2015 by 35 kilograms (kg) from the baseline year, whereas the grant provided 50 kg. In addition, losers of the lottery in these sites increased fertilizer use by 10 kg from the baseline year, consistent with the interpretation that a portion of the fertilizer windfall was shared with lottery losers. In sites with the seed treatment, winners of the fertilizer lottery increased fertilizer use by only 24 kg (31% less than in control villages) from the baseline year, while lottery losers increased fertilizer use by 12 kg (20% more than in control villages).

These estimates are consistent with the interpretation that there was greater social taxation of “windfall fertilizer” in treatment sites. One explanation for this greater social taxation is the seed treatment information campaign in treatment sites emphasized the importance of fertilizer to improving maize yields (especially when applied to hybrids). From this perspective, it is not surprising that fertilizer lottery losers in treatment sites would have put greater pressure on winners to share their windfall fertilizer than losers would have done in control sites. When we carry out the same statistical exercise on fertilizer used on maize, we obtain a similar picture. Given this configuration, it is unsurprising that households that benefited from both the seed treatment and the fertilizer lottery basically only benefited from the seed treatment, with the added benefit of fertilizer largely erased by this pattern of sharing. This pattern of differential social taxation across treatment and control villages is thus consistent with the pattern of estimated treatment effects shown in Table 3.[31]

---

[31] Additional directional evidence for this interpretation would be if increased fertilizer use was socially taxed most for lottery winners with high degree centrality to their social networks. We are unable to conduct that analysis, however, due to a lack of any network data in our seed control communities and only partial network data in our seed treatment communities from a study related to this work.

Table 7: Effects on total fertilizer use on farm (kg).

Control Households	
Baseline Mean	9.69
Midline	9.86**
	(3.48)
Endline	5.42
	(2.72)
Additional "Seed Treatment Only" Effect	
Baseline	3.07
	(2.85)
Midline	2.35
	(5.61)
Endline	-0.31
	(3.89)
Additional "Fertilizer Winner Only" Effect	
Baseline	-0.43
	(1.65)
Midline	25.46***
	(5.51)
Endline	5.10
	(3.64)
Additional "Seed Treatment & Fertilizer Winner" Effect	
Baseline	-1.03
	(3.41)
Midline	16.40**
	(5.03)
Endline	3.12
	(5.21)
Observations	1750
Matched Pair Fixed Effects	Included

All specifications include pair indicator variables as controls.

17 observations dropped with extreme values (greater than 500 kg).

Standard errors clustered at the village-level in parentheses.

* = 10% significance, ** = 5%, *** = 1%