Agriculture and Human Values

Female access to fertile land and other inputs in Zambia: Why women get lower yields

--Manuscript Draft—

This is a post peer-reviewed, pre-copy edited version of an article published in *Agriculture and Human Values*. The final authenticated version is/will be available online at:

DOI: 10.1007/s10460-018-9872-6

Female access to fertile land and other inputs in Zambia: Why women get lower yields

William J. Burke, Serena Li, Dingiswayo Banda

Accepted: 18 June 2018

Abstract: Throughout the developing world, it is a well-documented fact that women farmers tend to get lower yields than their male counterparts. Typically this is attributed to disproportionate access to high-quality inputs and labor, with some even arguing there could be a skills-gap stemming from unbalanced access to training and education. This article examines the gender-based yield gap in the context of Zambian maize producers. In addition to the usual drivers, we argue that Zambia's patriarchal and multi-tiered land distribution system could disfavor women with respect to accessing quality soils. We are uniquely able to control for soil characteristics using farm data from a sample of 1,573 fields with accompanying soil analysis. We find an expected difference in yields, but no evidence of a gap in unobserved characteristics, like skill, after controlling for access to inputs, especially quality soil, suggesting women are indeed disproportionately disadvantaged. We discuss how our findings could be used to develop self-targeting policy interventions that could empower women and would be consistent with the government's stated equity goals.

Keywords:
Gender yield gap; productivity; soil quality; sub-Saharan Africa; Zambia

Abbreviations:

CSO	Central Statistical Office (Zambia)
ha	Hectares
IAPRI	Indaba Agricultural Policy Research Institute
kg	Kilograms
MAL	Ministry of Agriculture and Livestock
OLS	Ordinary least squares
RALS	Rural Agricultural Livelihood Survey
SOM	Soil Organic Matter
USDA	United States Department of Agriculture
ZNFU	Zambia National Farmers Union

Author addresses:

William J. Burke (corresponding author), burkewj2@gmail.com

Agricultural and Food Policy Consulting, 3016 Saint Paul St., Baltimore, MD 21218, USA

Serena Li, serenali@stanford.edu

Freeman Spogli Institute for International Studies, Stanford University, 616 Serra St, Stanford, CA 94305 USA

Dingiswayo Banda, dingibanda@gmail.com

Ministry of Agriculture and Livestock, Mulungushi House, 2nd floor East, Lusaka, Lusaka, Zambia

William Burke, PhD, is an Agricultural Economist who has lived, worked and/or published research based in several African countries including The Gambia, Zambia, Malawi and Kenya. He has been an Assistant Professor, International Development, at Michigan State University's Department of Agricultural, Food, and Resource Economics and a Research Scholar at Stanford University's Center on Food Security and the Environment. Since September of 2016 Burke has been conducting research at Agricultural and Food Policy Consulting with several projects continuing in Southern Africa (https://www.afpconsulting-burke.com/). Burke has been honored with the 2009 award for best paper in *Agricultural Economics*, the flagship journal of the International Association of Agricultural Economists; The Horizon Ribbon Award, from Washington College; and the 2017 Bruce Gardner Memorial Prize for Applied Policy Analysis awarded by the Agricultural and Applied Economics Association.

Serena Li is a food and agriculture research consultant to Stanford University's Woods Institute for the Environment and ClimateWorks Foundation, a climate philanthropy with global impact. She received her graduate degree from Stanford, where she investigated issues of food security – in particular, the interplay between global food systems and the changing climate. Prior to Stanford, Li worked in Tanzania as a consultant for the Tanzania Horticultural Association, a USAID partner and leading advocate of the country's horticultural industry. Serena also has a decade of experience in the private sector. She has worked in communications and consulting for Fortune 500 companies, agencies, and start-ups in New York and San Francisco. Serena is from Maryland and received her B.S. from Northwestern University in Learning and Organizational Change with a minor in History.

Dingiswayo Banda is Chief Economist in the Policy and Planning Department of the Zambian Ministry of Agriculture and Livestock based in Lusaka, Zambia. He earned his Masters Degree from Michigan State University's Department of Agricultural, Food, and Resource Economics. The Policy and Planning Department has a role of ensuring that policies and objectives of the Ministry are maintained and coordinated in an effective manner on a continuous basis as well as reviewing these policies and objectives and making necessary changes to reflect the current demands of the agricultural sector. The Department also has the mandate for overall Policy formulation coordination, programme planning, monitoring and evaluation, budgeting, donor co-ordination and mainstreaming gender and HIV/AIDS and other cross cutting issues within ministerial programmes.

Acknowledgements:
Our foremost gratitude is to the participants in the 2012 Rural Agricultural Livelihoods Survey and the key informants at the Zambia Agricultural Research Institute, the Ministry of Agriculture and Livestock, the Ministry of Gender and Child Development, and focus group participants during the spring of 2015. We thank IAPRI for collecting and sharing the data used for this study. We thank Roz Naylor, Danielle Nierenberg, Wally Falcon, Agnes Quisumbing, two anonymous reviewers, the editor and participants of a special session of the 2016 Agricultural and Applied Economics Association annual meeting for useful comments.

Compliance with Ethical Standards:
All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study. The authors declare that they have no conflict of interest.

Female access to fertile land and other inputs in Zambia: Why women get lower yields

Introduction

In much of the developing world, male farmers get higher yields than female farmers. This is well-documented by several studies (Udry 1996; Adesina and Djato 1997; Tiruneh et al. 2001; Gilbert et al. 2002; Horrell and Krishnan 2007; Larson et al. 2015; Kilic et al. 2015). The *reasons* women tend to get lower yields, on the other hand, continues to be the focus of empirical work and policy discussions.

The gender yield gap is certainly an important, attention-worthy matter. For one, most farmers in Zambia (the focus country of this analysis) are women. According to the farmers' union, 52% of Zambians working in agriculture are women (ZNFU 2013), but women obtain 20-25% less output per hectare (this study). Very recent literature highlights reducing gender-based differences in agricultural productivity as a key opportunity for stimulating smallholder agricultural growth more generally, both now and for coming generations (Kilic et al. 2015).

One argument as to why men get higher yields is that men are innately better farmers, although there is no empirical evidence showing a clear "ability gap" between male and female farmers that we are aware of. Alternatively, the gender yield gap could arise from disparate access to productive resources, including land, labor, extension, education, or input markets. Evidence of the resource gap argument has been supported by numerous empirical studies (Udry 1996; Adesina and Djato 1997; Goldstein and Udry 1999; Tiruneh et al. 2001; Gilbert et al. 2002; Horrell and Krishnan 2007; Aguilar et al. 2015; Karamba and Winters 2015; Oseni et al. 2015).

Unfortunately "ability" is notoriously difficult to measure. With this analysis we consider the statistical implications of omitted variables, whether they be unobservable (like ability) or just left out when estimating yield determinants, and specifically the bias one could expect omitted variables to cause if they are correlated with gender. We show that disparate access to factors of production, and especially including soil quality, explains much if not all of the difference in yield between genders. We show that, while it is not possible to demonstrate superior ability amongst men, it is possible to demonstrate there is no evidence of an ability gap between male and female Zambian maize farmers. The primary contribution of this study is our consideration of differences in soil quality in plots managed by men and women as a driver of yield differentials. In addition to the farm management and factors of production previously considered, we explicitly control for soil quality from samples collected from the fields used in our study. We use data from 1,573 maize fields in a nationally representative Zambian sample. Unlike previous studies we also model the allocation of land between men and women, and show that the allocation of more fertile plots is biased against women. This suggests an under-explored source of the productivity gap which traces its roots to the culturally-entrenched land allocation systems that are biased against women.

Our objective is to describe the yield gap between men and women amongst Zambian maize farmers and identify the key drivers of the differences in land productivity. To meet the objective we answer three questions:

1) What is the difference in maize yields between male and female controlled fields?
2) What happens to our simple estimate of the gender-based yield gap if we control for other factors of production, including soil quality?
3) Is there direct evidence that female farmers are assigned fields with lower soil quality?

The disenfranchisement and lack of equality for women is explicitly recognized as a problem in the majority of African nations (e.g., Alozie and Akpan-Obong 2017). It is understood that a more level playing field between and among genders would be beneficial for all of society, not just for women (e.g., Kilic et al. 2015). Reducing gender-based disadvantages in agriculture, or any part of society, has the potential to unlock or make scalable tools for innovation and improving productivity that might otherwise remain underdeveloped. Systemically disadvantaging half of the farming population of an agrarian society while trying to promote economic development is like trying to farm with one hand tied behind your back. This study investigates whether soil quality, one of the most important determinants of agricultural potential, is allocated in a way that systemically disfavors women. To our knowledge, no similar analysis of soil and inequality that is nationally representative while considering this scope of determinants exists.

Background

The gender-based yield gap

Several studies across sub-Saharan Africa find gender yield gaps ranging from 11% to 40%. Most frequently, analysts have attributed the gap to differences in access to labor and non-land inputs like seed and fertilizer (Udry 1996; Adesina and Djato 1997; Tiruneh et al. 2001). Some studies do assert a portion of the gap is attributable to differences in skill, citing a disproportionate level of access to extension advice or "experience" (Tiruneh et al. 2001; Gilbert et al. 2002; Horrell and Krishnan 2007; Larson et al. 2015).

These studies do not, however, consider differences in the quality of soils being farmed by women. There have been a few quantitative studies across the developing world investigating the gender-soil relationship, but data are imprecise or inconsistent. In Burkina Faso (Udry 1996) and China (de Brauw et al. 2008) researchers have used self-reported soil quality and find little or no quality difference by gender. Fuentes and Wiig (2009) examine soil salinization and erosion amongst Peruvian farmers and find they make no difference for yields by gender. Goldstein and Udry (1999) empirically measure soil characteristics similar to those we examine when investigating gender differences. In addition to self-reported data, they measured soil organic matter and acidity, finding that female farmers have significantly less organic matter in their soil, but find no difference in females' soil acidity. Our study builds on this literature focusing on Zambia, where land tenure and plot allocation are complex issues.

Land allocation in Zambia

There are several stages in land allocation in Zambia, and thus several ways women could potentially be disadvantaged. To understand the allocation process, though, it is worth briefly describing Zambian land tenure institutions.

In Zambia all land is either state land or customary land and allocated under a bifurcated system rooted in British colonialism. State land is protected under the jurisdiction of the government, is legally titled, and can be bought or sold. Customary land is managed via the discretion of regional chiefs (Honig and Mulenga 2015). A third, pseudo-category of land is customary land that has been titled and converted to “leasehold” land. Legally the rights to this land are currently the same as customary land, though a draft of a constitutional revision would treat it as state land.

Following independence in 1967, it was estimated that 6% of Zambian land was state land and the remaining 94% was customary. While it is unclear how these estimates were calculated, these figures have remained in public consciousness, despite informal agreement that the numbers have changed. Recent studies suggest that there has been a significant shift, now with customary land constituting just 60% of the country (Honig and Mulenga 2015). With less land available and more land contested, land tenure in the agricultural sector has become more complicated, and potentially more discriminatory.

Customary land is allocated through a multi-tiered system, and one best described by the literature dedicated specifically to the subject (see also Smith and Naylor 2014; Sitko et al. 2014). In short, tribal chiefs (usually men) divide customary land among village headmen and (less often) headwomen. Village heads then disperse the land to male and female household heads, which then allocate land among male and female household members. For our purposes, it is sufficient to understand that most land is still customary and Zambia’s patriarchal history *could* disadvantage women with respect to allocation of quality land.

Our study examines productivity at the household level. Household heads may be male or female; typically female heads of households are either divorced or widowed. Within a household, there are males and females; female adults are usually the wives of the male head of household. Widowed or divorced women could lose access to productive land because in Zambia, and many other African societies, women are more likely to have married into a man’s family compared to the other way around. Staple crops, which also usually bring in the most income, are grown on the main family plot. In Zambia, this is overwhelmingly maize because of heavy government subsidies for the crop (Jayne et al. 2018).

The importance of soil

Once land is assigned to a household, it is can be passed down through generations, and households are largely beholden to the quality of soil endemic to their village. The soil in Zambia tends to be acidic, especially in the north, acidic and sandy in the west, and sandy loam/clayish in the east and center of the country. While these trends are indicative, there can be important variations in local soil quality, even within villages (Burke et al. 2016).

All plants require nutrients, and 14 of the 17 essential elements required for plant growth come from soil. Soil's physical, chemical, and biological characteristics largely determine the presence and availability of these elements. Physical properties include soil particle size (ranging from large sands to small clays). Generally, more clayish soils hold more moisture and nutrients, and are more productive. Chemical soil properties include soil pH level (where pH<7 is acidic), which influences root growth and nutrient availability. On acidic soils root growth is stunted and the bonds that hold essential nutrients are stronger, meaning nutrients can be unavailable to plants, even if they are present. Farming thus tends to be much less productive on acidic soils. Biological soil organic matter also indicates nutrient availability and whether nutrients are likely to be available to plants. Soils tend to be more productive if they have more organic matter. In short, the properties of soil are collectively key in determining both the *presence* and the *availability* of nutrients that are critical for healthy crop production (Jones et al. 2013).

Data

Our analysis uses data collected by the Indaba Agricultural Policy Research Institute (IAPRI) in collaboration with the Zambian Ministry of Agriculture and Livestock (MAL) and Central Statistical Office (CSO) as part of the 2012 Rural Agricultural Livelihood Survey (RALS). A nationally representative sample of 1,573 households was asked detailed questions spanning demographic, economic, and social categories.

The data are described in detail in Burke et al. (2016). Of an original sample of 1,643 fields, we exclude 70 (<5%) potential outliers for the continuous variables of yield (less than 57 or more than 17,000 kg/ha), fertilizer application rate (more than 2,357 kg/ha) and seed application rate (more than 271 kg/ha). In each case the cutoff point is determined by the top or bottom 1% of the overall sample distribution. It is worth noting that our conclusions are robust to including these potentially influential cases, but we prefer to report the results using the less noisy data (with potential outliers excluded). Other variables are categorical and thus cannot contain outliers.

In addition to the unique soil data, we highlight the availability of a few other key pieces of information. For one, the RALS survey asked detailed questions regarding field management.

There are data on the actions taken by farmers, and importantly for this study, the gender of the person deciding how fields would be managed.[1] There are data on the availability of inputs, age, education, access to extension and so on, all of which could determine yield variations. Some of the responses most interesting to us include the gender of the primary decision maker and other demographics, and agricultural practices such as the type, quantity and timing of various inputs actually used, including fertilizer, seed, herbicides, and pesticides.

<<Tables 1 & 2 about here>>

In addition to the RALS farm management and demographic data, we use soil characteristics taken from the fields used in this analysis, which are the households' largest maize fields. The soil samples were analyzed at the Zambia Agricultural Research Institute (ZARI), the research and extension branch of MAL. Measured soil characteristics are acidity (pH), texture (categorized using the USDA "soil triangle" and ranked by clay content), and soil organic matter (SOM). These indicators represent the soil's chemistry, physics and biology respectively, which, in combination, can describe a soil's fertility. Means for the variables used in this analysis are disaggregated by gender groups and presented in Table 1. Table 2 shows more general distribution statistics for the full sample.

Conceptual Framework

To answer our first research question we will begin with a simple linear Ordinary Least Squares (OLS) regression of maize yield on a binary variable indicating whether the field manager is female. In this regression the constant term will be equal to the sample mean yield from male-managed fields and the coefficient on the female indicator variable is the difference in average yields between male and female managed fields. A negative coefficient estimate in this simple regression would indicate that mean yields are lower for women.

[1] To be clear, the data regarding who manages the largest maize field (where the soil samples were taken after the 2012 harvest) were asked with respect to the 2011 harvest. For lack of a better option, we assume the manager of the largest maize field did not change from 2011 to 2012.

Under the standard assumptions required to interpret a negative coefficient as a causal relationship (i.e. to assume that yields are lower *because* the farmer is a female, and *only* because a farmer is female), we would be assuming there is some systemic difference in the processes that contribute to yield realizations between men and women. We argue this would be a baseless assumption, and so we look further into what may be driving the statistical gender yield gap.

Why do female farmers have lower yields?

Forgoing the possibility that random forces systematically alter the factors that determine yields for women, a negative difference between yields of women and men could have a few possible causes. Specifically, we believe the negative coefficient is driven by an "omitted variable problem".[2] To explain how omitted variables could be influencing the estimated "effect" of being female on yields, it is helpful to discuss the aforementioned bivariate model and OLS estimator:

$$yield_i = \alpha + \beta_1 female_i + u_i \qquad (1)$$

where *yield* is the kilograms of maize harvested per hectare for field *i*, and *female* is the indicator equal to one if the field is managed by a female, and 0 otherwise. When we estimate these parameters using OLS, the results are tantamount to some of the basic statistics in Table 1. Since other variables are omitted, however, it is unlikely that we should interpret our estimates (particularly of β_1) as causal – the estimates are biased unless *u* and *female* are uncorrelated.

There are important consequences if *u* and *female* are correlated. Suppose, for example, soil quality (*quality*), which certainly affects yield, is correlated with *female* and is unobserved

[2] Strictly speaking, this is "endogeneity" which occurs in multivariate statistics any time there is correlation between the unobserved residual term and explanatory variables. One form of endogeneity occurs when there is bi-directional causality or "simultaneity" between the dependent and explanatory variables. There is almost certainly no such endogeneity in our model since all of the explanatory variables are pre-determined. Another reason there may be endogeneity is if the unobserved residual includes omitted relevant variables that are also correlated with explanatory variables. This is the type of endogeneity around which we focus our discussion. See Wooldridge (2002) for a detailed treatment of the subject.

(or not included) in equation 1. In notation form, we can say $u_i = \beta_2\, quality_i + \varepsilon_i$, where ε_j is the new error term. Now the model becomes:

$$yield_i = \alpha + \beta_1 female_i + \beta_2 quality_i + \varepsilon_i \qquad (2)$$

and β_2 is the correlation between soil quality and yield. How we specifically measure *quality* will come later, but for now it is sufficient to say we expect higher yields on higher quality fields (i.e., β_2 is positive). Further suppose there is some relationship between *female* and *quality*, such that:

$$quality_i = \gamma + \delta_1 female_i + v_i \qquad (3)$$

We can state equation 3 without any controversy because we have not made claims about either the causal relationship between *female* and *quality* or the value of δ_1. If equations 2 and 3 hold, a straightforward algebraic simplification can show the estimated coefficient in equation 1 (call this β_1^{eq1}) is biased, and specifically $E(\beta_1^{eq1}) = \beta_1 + \beta_2\delta_1$. So, if soil quality improves yield ($\beta_2 > 0$, a certainty by definition) and females are disproportionately assigned to lower quality fields ($\delta_1 < 0$), the omitted variable bias in equation 1 would usually lead to a negative estimate of β_1^{eq1}, *even if being managed by a female farmer has no direct effect on yields*. This could lead us to conclude there is a negative effect of being female on yield where none truly exists. The real correlation, in this case, would stem from the interaction of the two relationships: the correlation between *quality* and *yield*, and that between *quality* and *female*.

Alternatively, to argue women are simply not as good at farming is tantamount to arguing the omitted variable is the farmer's innate ability. By this logic, ability is positively correlated with yield and negatively correlated with femaleness, which would cause us to expect a negative coefficient on the female indicator variable. When innate ability and soil quality are both omitted (along with everything else), we cannot make an empirical argument for why our estimate of β_1 is positive or negative if we expect it to be zero.

In either case, we might expect the coefficient estimate to change if we are also able to control for omitted variables. Controlling for *quality* (and other determining factors) can be done explicitly using our data. Ability, on the other hand, might be controlled for using "proxy variables." Since ability is not something that has a specific metric, we rely on things that are measurable and which we expect to be correlated with ability. For example, we may use the years of formal education, the farmer's age, or whether the farmer has received advice from an extension agent as indications of ability.[3]

In order to unpack the *reasons* female farmers tend to have lower yields, in this analysis we will incrementally add the determinants of yield to the model in equation 1. In a fundamental agronomic sense, the only thing that determines production of plant matter is the availability of just a few (seventeen) essential nutrients (Jones et al., 2013). We discussed earlier the role that soil plays in determining the availability of those nutrients, and obviously, fertilizer can play an important role if the soil (regardless of its general characteristics) are lacking those nutrients. Similarly, variables such as the seed rate and the type of seed used obviously belong in the yield equation – more seeds can be expected to produce more plants and grain per hectare (at least within the range of seed rates found in our sample), and some varieties will more effectively convert available nutrients into crop output.

We also control for "other management" practices, which will include early planting, the use of agroforestry, number of weedings and other crop mixes. Early planting matters because, particularly during the beginning of the rainy season, the availability of mobile nutrients like Nitrogen can drop off quickly over time. Agroforestry refers to the use of trees that can serve useful purposes such as fixing atmospheric nitrogen or providing mulch (fallen leaves) to maintain soil moisture and add nutrients. Certain crop mixes (especially legumes) also fix atmospheric nitrogen and can increase yields of nearby crops, whereas a lack of weeding could mean field crops must compete with weeds for available nutrients (See Jones et al. (2013) for a general overview of yield determinants, Kilic et al. (2015) for a discussion of issues in sub-Saharan Africa, or Burke et al. (2016) for a discussion of matters important in Zambia specifically).

[3] An important side note, for this purpose, is that it does not matter *how* these observable proxies are correlated with "ability", just that they *are* correlated. For example, we might expect that a farmer who seeks extension advice will be trained and then have greater ability. On the other hand, perhaps skilled farmers don't need extension advice, in which case those who visit extension workers would have less ability. Either way, as long as there is systemic correlation, controlling for advice seeking is a suitable proxy for ability. The importance of ambiguous correlation would be different if we were trying to measure a causal relationship between yield and receiving advice from an extension worker.

We use “education” variables as the aforementioned proxies to control, to the extent possible, for a farmer’s ability. These variables are meant to capture the variations in farm management that are either not explicitly included in the discussion so far, or are otherwise not observed (for example, applying fertilizer opportunistically in areas where it will be most effective). We incorporate provincial fixed effects as a proxy for net combined effects of differences in weather, infrastructure, the quality of public goods, and other omitted factors that vary across provinces)

Again, we will add these variables incrementally starting with provincial fixed effects, then input use, other farm management, education, and soil characteristics (in that order) to examine changes in the “effect” of a field being female-controlled. As we add determinants that are correlated with being female, we expect the estimated “female effect” to change – if the estimated effect does not change it means the added variables are not systemically correlated with being female in our data.

Primarily the incremental revisions to the model will be simply adding explanatory variables to a linear model, as in going from equation 1 to equation 2. When incorporating soil quality variables, however, we will follow the prevailing recent trend in the literature and estimate a threshold-based quasi-linear model (Marenya and Barrett 2009a; Marenya and Barrett 2009b; Matsumoto and Yamano 2013; Burke et al. 2016; Burke et al. 2017). The estimation procedure is briefly described in Appendix A. In short, the models including soil quality will produce more than one estimate of the female “effect” (differing according to soil quality regimes).

Finally, we will directly examine whether female farmers are systemically disfavored with respect to soil quality when fields are allocated by examining how being female relates to the probability of being allocated poorer quality soils using some descriptive and multivariate analyses.

Results

Examining the “female effect”

A bivariate comparison is presented in model (i) of Table 3a, and shows that the average male managed farm produced about 2.4 metric tonnes (mt) of maize per cultivated hectare. By contrast, the average female-managed hectare produced nearly 0.6 fewer mt, or 23% less. This difference is statistically significant at the 1% level and within the range of gender-based yield gaps shown in other literature. To be clear, we believe this is a biased estimate of the “effect” of female management because there are omitted relevant variables. When we add provincial fixed effects (i.e., a binary variable for each province in Zambia, Table 3a, model ii) they are jointly significant at the 1% level, showing there are meaningful differences in mean yield between provinces, but the difference between male- and female-managed fields is still estimated as nearly 0.5 mt. So, while both province and female headedness are highly correlated with yield, the relatively small change in the female coefficient suggests they are not highly correlated with each other.

When we add farm management variables the story changes dramatically. In model (iii) we add fertilizer application rate (kg/ha), seed application rate (kg/ha) and a binary indicator variable for whether the seeds are for hybrid or OPV (usually jointly called “improved”) plant varieties. Unsurprisingly, seed and fertilizer rates are highly significant (and since seed rate is never zero, the intercept no longer has a very meaningful interpretation by itself). Improved seed use has a positive coefficient, as expected, but is not statistically significantly different from zero. When we add seed and fertilizer to the regression the “effect” of female management is reduced by more than half. In other words, when using the same seed and fertilizer, the difference between male- and female-managed fields is about 220 kg/ha (significant at the 5% level), which is less than 40% of the naïvely-estimated difference in model (i). Together, these results show that female management is negatively correlated with improved seed use and seed and fertilizer application rates.

<<Tables 3a & 3b about here>>

The results described for model (iii) are remarkably robust as we add variables for other management and proxies for ability (models (iv) and (v)). “Other management” includes early planting, use of agroforestry for soil maintenance, the number of weedings and crop mixing, which are jointly significant at the 5% level. Three binary indicators represent “education” for whether primary, secondary, or post-secondary (including college or technical school) have been completed. We also include a dummy variable for whether the household received extension advice on conservation farming with the education controls. Interestingly, other management variables are jointly significant at the 5% level, but the ability proxy variables are not jointly significant.

Moreover, it is worth noting that including “other management” has virtually no impact on the female coefficient. So, while fertilizer and seed use (financially burdensome inputs) are negatively correlated with female management, other best practices (more skills-related activities) do not appear to be correlated to female management. Finally, in Table 3b we add soil quality variables in models (vi) through (ix), each of which fills two columns, as they are quasi-linear.

The first two models that include soil (models (vi) and (vii), Table 3b) both estimate a structural bifurcation of the yield function at a soil organic matter content of 1.3%, which is consistent with what one would expect based on agronomic literature (Jones et al. 2013). A Chow test suggests the structural equations are statistically significantly different from each other at the 1% level. Model (vii) includes both SOM and clay content variables, with clay effects that are also consistent with prior expectations, especially on the high SOM fields.

The last two models include soil pH instead of SOM and both estimate a structural break in the yield function at a pH of 5.6. Again, this is almost exactly where one would expect a structural break (ZARI 2002; Jones et al. 2013), and again, the statistical test strongly supports the significance of a structural break. Like model (vii), model (ix) in Table 3b also includes clay content variables, with estimated clay effects that are consistent with prior expectations, especially on the low pH fields. It is also worth noting that fertilizers in particular appear considerably less effective on lower quality soils.

<<Figure 1 about here>>

Most importantly for this study, however, the "effect" of female management, while still negative, has even less magnitude after controlling for soil quality (as little as 26% of the initial estimate) and, even more tellingly, is *not* statistically significant at any meaningful level in all but one regime of one of the models that include soil quality. The impacts of adding explanatory variables to our models on the "female effect" shown in Tables 3a and 3b are visually summarized in Figure 1.

Female access to fertile soils

The fact that controlling for soil characteristics renders the "female effect" on yields not statistically significant is compelling evidence that women have less access to quality land in our sample.[4] To further investigate whether fertile land distribution between the genders is unbalanced (our third question), we can observe some illustrative descriptive statistics, such as that 58% of male-managed fields have a high (more productive) clay content (> 45% clay) compared to 54% of the fields managed by female heads of household and just 37% of the fields managed by female members of male-headed households, or that only 17% of female-managed fields in male-headed households are in the (more productive) high pH range (> 5.6), compared to 29% of the male household head-managed fields.

Neither our earlier results nor the descriptive statistics, however, are a direct test of the hypothesis that women are disproportionately disadvantaged – which can be done with regression analyses. Unfortunately, it has proven difficult with these data to draw very nuanced conclusions on the correlations between female management and soil quality using multivariate methods (i.e., running regressions of soil quality variables on *female*, or *female* interacted with conditions like education and relationship to the household head, all while controlling for other

---

[4] It is worth noting that our preliminary, broader analysis did examine the possibility that the decrease in statistical significance of the "female effect" was simply a statistical artifact (i.e., whether the impact of adding soil on the "female effect" is driven by a degrees of freedom constraint because soil is added last in our analysis). We do not believe this is the case for two reasons. First, the subsets of our data that represent each soil regime contain hundreds of observations, and other important variables (e.g., fertilizer) are still significant after controlling for soil. Second, we find effectively the same results when we add soil variables *before* controlling for the nine variables encompassed by "other management" and "education". In other words, the decreased significance of the "female effect" is almost certainly not just a statistical artifact.

relevant factors) because the data quickly become too thin to meaningfully estimate a broadly specified model.

Caveats notwithstanding, we can demonstrate some interesting results and trends in more parsimonious models, including important variability in the disproportionate access to quality land. To this end, this section examines the relationships between soil characteristics and gender within the contexts of the other explanatory variables in Tables 3a and 3b. Again, since the models in this section necessarily omit relevant factors, the results should be thought of as interesting correlations, while any causal interpretation would require supposition.

<<Table 4 about here>>

Table 4 shows the results of regressions of pH on indicator variables for whether a female head of household or a female spouse of male head of household is managing the field.[5] This separation of female managers according to their relationship with the head is motivated by literature indicating its potential importance (e.g., Mason, et al., 2015). The first column shows the intercepts from these regression, which are actually the mean pH value on the fields that are managed by men in our sample. The coefficients on the indicator variables, then, are the difference in the means between the indicated group and men. Recall that lower pH values indicate more acidic and generally less productive soils. We see the mean pH on male-managed fields is about 5.4 in the Central province, compared to a mean pH that is 0.11 lower (or about 5.3) on fields managed by female heads of household, and 0.23 lower (or about 5.2) on fields managed by female spouses of male heads. Also note the difference between male-managed fields and those managed by female spouses of male heads in Central province is statistically significantly different from zero at the 1% level. In fact, we see that in 7 of the 10 provinces, the mean pH on fields managed by female heads of household are lower than the pH on fields managed by men – the exceptions being the Copperbelt, Muchinga and Northwestern provinces. Notably, though, none of these differences (negative or positive) is statistically significant at any meaningful level.

[5] In fact, 3 (or <5%) of the women we are calling "female spouses of male heads of household" are identified as "other female household member". Since 3 observations are too few for any meaningful analysis, we include these women in the group of spouses.

Similarly, in 7 of the 10 provinces we find the mean pH on fields managed by female spouses of the heads of household are lower – the exceptions in this case being the Luapula, Lusaka and Northern provinces. Unlike the differences for female heads of household, however, we find 4 of the 7 negative differences for female spouses are significantly different from zero at the 10% or 1% level. The biggest discrepancies are in the Copperbelt and Muchinga provinces, where pH on fields managed by female spouses are roughly 0.5 lower than the pH on fields managed by men in the same province.

Interestingly, the female spouses of male heads in Lusaka province are farming on land where pH is significantly *higher* than the land of their male counterparts. It is also worth noting, however, that all of the gender groups in this analysis in Lusaka have a "good" pH that is, on average, above the critical threshold of 5.6 that was estimated using these data.

<<Table 5 about here>>

Table 5 presents analysis similar to that in Table 4, but in Table 5 the dependent variable is SOM. Here again, the disproportionate distribution of quality soil characteristics is more evident amongst female field managers who are not household heads. Specifically, 6 of 10 provinces show lower SOM on fields managed by female spouses compared to the SOM of male-managed fields. Three of the 6 differences are statistically significantly different from zero at the 10% level or lower (plus Muchinga, where the difference is significant at the 12% level). Most notably the SOM on fields managed by female spouses of male heads is nearly a full percentage point lower (a difference significant at the 1% level) than the SOM on male-managed fields. Here too, though, there are anomalies. For example, results in Central province suggest female spouses have significantly *higher* SOM on their fields as compared to the fields of men in the province. Again, however, it is worth noting that male-managed fields in Central have about 1.9% SOM on average, which is well above the "high SOM" threshold.

Finally, we examine the relationship between gender and soil by estimating models of pH and SOM as a function of gender (female heads and spouses combined into one group) in the context of years of education. That is, the explanatory variables are indicator variables for each year of education as well as an interaction term for the female indicator. The results of this model can be summarized as in Figures 2 and 3, where we show the expected value of pH and SOM respectively for men and women over the number of years of formal education (note, we truncate these results at 10 years of education, because the estimates are thin beyond this point, especially for female-managed fields). Expected values for female (male) managed fields are shown as a dashed (solid) line – 90% confidence intervals are shown as light grey lines.

Figure 2 shows roughly similar expected pH levels among male and female farmers of lower education levels, up to around 5 years of formal education. Amongst more educated farmers, it seems that women are disproportionately likely to be allocated more acidic (less productive) land, where even the 90% confidence intervals do not overlap between 6 and 8 years of education. Figure 3 shows a similar disparity among more educated male and female farmers, but the difference is only statistically significant at 7 years of formal education.

<<Figures 2 & 3 about here>>

In summary, we find the sum of evidence indeed suggests female field managers are disproportionately likely to be allocated less productive land, which is consistent with our previous findings that the "female effect" is statistically nil after controlling for soil quality. However, when we use regression analysis to identify direct systemic relationships between gender and soil quality we find important variability in the disadvantages facing women. First, it seems that female spouses of male household heads are even more likely to be allocated poorer quality soils than female heads of household. Moreover, some geographic locations, such as Eastern province, do not show statistically significant differences in the distribution of quality soil between gender groups, while others, such as Northwestern province, show disadvantages for women across soil fertility measurements. Interestingly the disproportionate access to quality soil seems more pronounced amongst more educated Zambian farmers. In all cases, however, we must consider our results under the caveat that we do not have enough data to simultaneously estimate all parameters in a comprehensive model of land access. As such, we cannot make very definitive empirically motivated statements about the underlying mechanisms of these correlations. That women are disproportionately allocated fields with poorer quality soil, in general, seems quite clear. The problem we face when we begin to disaggregate groups of female managers, as we have, by education, location or relationship to the household head, is that the estimates become thin for all but the most frugal models. The only way to resolve this problem or to incorporate even more characteristics (e.g., tribe, relationship to village leaders) in future analyses of this type will be through more data collection.

Discussion and conclusion

In this study we used data from a sample of 1,573 Zambian maize fields with accompanying soil analysis to answer three questions. First, what is the difference in maize yields between male and female controlled fields? Using a simple comparison, we found female farmers tend to harvest 23% less per hectare than their male counterparts. We note, however, this is a naïve observation and, pointing out the role of omitted variable bias, we emphasize it would be a mistake to interpret the correlation between gender and yield as a causal relationship

Second we ask what happens to our simple estimate of the gender-based yield gap if we control for other factors of production, including soil quality. We see the estimated difference in yields between the genders decreases in both meaningful and statistical significance when we begin to control for the factors of production. In fact, after controlling for soil quality and other factors, we find no statistically significant difference between the yields on male- and female-managed fields – In other words, using our least biased estimator we conclude female management has no independent effect on yields.

Based on these results, we conclude the gender-based yield gap is almost entirely the result of disparate access to productive inputs, including seed, fertilizer and more productive soils, and not the result of other systemic differences in unobservable factors between the genders. One might argue women use less or lower quality inputs because they are less able farmers, but such an argument is not supported by our evidence – proxy variables for unobservable factors, like "ability", show no evidence of being correlated with the gender of the field manager, and including these proxies has no impact on the estimated "female effect". Furthermore, while including financially burdensome inputs like fertilizer and seed has a large effect on the female coefficient, we find no similar effect when including more obviously skill- or ability-related management variables. We thus contend the difference in fertilizer and seed use is driven primarily by differences in access, not differences in some underlying skill or ability as a farmer. Explicitly testing the hypothesis that African women have less access to productive inputs like fertilizer and seed is not within the scope of this study, but has been extensively addressed (and is supported) in many prior studies (e.g., Adesina and Djato 1997; Tiruneh et al. 2001; Gilbert et al. 2002; Horrell and Krishnan 2007).

Our third question is whether there is direct evidence that female farmers are allocated fields with lower soil quality. The fact that controlling for soil characteristics renders the "female effect" on yields statistically insignificant is compelling evidence that women indeed have less access to quality land in our sample. In seeking more direct evidence, we examine some descriptive figures and estimate parsimonious multivariate models. Consistent with earlier findings, we see women are more likely to be allocated poorer quality soils, and the difference is more prevalent among female spouses of male household heads compared to female heads of household. Moreover, we find significant geographic variations in female access to higher quality soils. For example, data from the Eastern province shows no statistically significant differences in the distribution of quality soil between gender groups, while others, such as Northwestern province, show disadvantages for women for soil fertility measurements like pH and soil organic matter. Interestingly the disproportionate access to quality soil seems more pronounced amongst more educated farmers. In all cases, however, we must consider our results under the caveat that we can only estimate narrowly focused models using these data for answering our third question. That women are disproportionately allocated fields with poorer quality soil, in general, seems quite clear. The only way to put a finer point on our understanding of the mechanisms of these disadvantages, or to incorporate more characteristics simultaneously in future analyses of this type, will be through more data collection.

We expect these findings to be meaningful to policy makers in African countries. The culturally entrenched disadvantages faced by women in many parts of Africa have a long history (Falola and Amponsah 2012). The disenfranchisement of women is explicitly recognized as a problem in the majority of African nations (36, including Zambia) that have signed and ratified the "Protocol to the African Charter on Human and Peoples' Rights on the Rights of Women in Africa" (the "Maputo Protocol"), followed by the subsequent declaration of the 2010's as the "Decade of Women".[6]

It is understood that a more level playing field between genders would be beneficial for all of society, not just for women. The Maputo Protocol says nations are "firmly convinced that

[6] In a recent special issue of *Development Policy Review*, Alozie and Akpan-Obong (2017) provide a useful summary of these two agreements.

any practice that hinders or endangers the normal growth and affects the physical and psychological development of women and girls should be condemned and eliminated" and they will "promote women's access to and control over productive resources such as land." The Zambian government itself has indicated a commitment to female empowerment by the creation of the Ministry of Gender and Child Development in 2011 and the National Gender Policy of 2014.

There is clearly a desire and need for policy to do more. For one, access to quality land and other inputs by women could be a more deliberate policy in deeds, not just words. In this context we must discuss agricultural input subsidies because many African countries spend considerable resources each year on them – input subsidies are the primary policy tool used to promote productivity and reduce poverty (Jayne et al. 2018). These subsidies might be an example of how women could be deliberately targeted, but so far this has not been the case. We believe subsidies have not effectively targeted women for several reasons. Most obviously, subsidy targeting has focused on increasing national production, not decreasing input access disparities (gender-based or otherwise). As a result, most subsidy beneficiaries tend to be relatively well-off and imbalances are exacerbated, not mitigated, by subsidy programs.

Second, while targeting subsidies could be an intervention point to increasing women's access to inputs in theory, the evidence shows effectively targeting specific households (where it has been attempted) is notoriously difficult (Jayne et al. 2018). Targeting individual members within a household will be even more difficult. The specific challenge in this case is that ostensive goals of policy makers – the fair treatment and improved livelihoods of female farmers – is in contrast with many cultural norms of the people tasked with implementing policies – an historically patriarchal system of allocating resources. Targeting goals could be unrealistic when individual beneficiaries are identified "on the ground" by patriarchal small groups or individuals, or when "targeted women" return to a male-headed village or household where inputs can be confiscated.

Addressing gender inequality in African agriculture will require a more thoughtful approach and going beyond "targeting women" with agricultural subsidies.

An altogether different policy design in both the near and long term may be needed to genuinely address unfairness based on gender. In the near term, it may be more effective to target the *disadvantages* than it is to target the disadvantaged *people*. For example, suppose a program could be used to improve soil quality where it is needed most. Focusing on interventions to improve the lowest quality soils would benefit the most disadvantaged farmers by definition – this will include men, but if our findings are correct the benefits are likely to disproportionately benefit women regardless of any gender-based targeting criteria. Admittedly, this argument sounds very similar to the argument for fertilizer subsidies (increasing access to fertilizer for everyone should be most advantageous to those with the least access). The key difference is that nutrient deficiencies addressed by fertilizers are ubiquitous and relatively evenly distributed – almost everyone has a reason to want fertilizer. This is why we can expect the benefits of increasing access to fertilizer will be concentrated among those with the most social influence, which is very frequently the outcome. Conversely, soil *quality* deficiencies are disproportionately a problem faced by women, so there is less vulnerability to elite capture of benefits (though it's certainly still a risk). Of course minimizing nutrient deficiencies is an important objective and we are not advocating that it be ignored, but maximizing soil quality, we think, is an opportunity to address an agricultural problem (low yields) and a social problem (gender inequality) at the same time and with realistic policy objectives.

Notably, promoting healthy soil management through farmer extension services is remarkably affordable compared to current government budgets for agricultural subsidy spending. One way to promote soil quality is to provide extension handbooks or pamphlets on soil management. Such handbooks have been written (e.g., ZARI 2002), but there are effectively none actually in circulation because there is no budget to print them. Better still would be to fund the extension agencies that are effectively nonoperational due to budget constraints. Of course, the effectiveness of pamphlets, handbooks or extension officers is dependent on factors like literacy and access to training programs. Even in the most remote parts of Africa, technological revolutions in access to phone networks and the internet have and will continue to make disseminating information more affordable than ever, but this is only helpful to people who can understand it, which brings us to the second point in the paradigm adjustment.

In addition to short-term efforts to address immediate disadvantages, improving the livelihoods of women in agriculture is a goal that requires long-term efforts to address the historical norms that have brought about their disadvantages. Improving the livelihoods of women in agriculture cannot be separated from, for example, improving girls' access to education, women's literacy, women's access to agricultural training and, very importantly, the protection of private property.

It may be that efforts since the Maputo Protocol have made agriculture more equitable, but it is certain that substantial room to improve remains open. The results of this study suggest that if countries intend to honor their commitments to enabling female farmers – if the "Decade of Women" in Africa is to be more than rhetoric – the explicit and implicit disadvantages women face require a high priority.

References

Adesina, A.A., and K.K. Djato. 1997. Relative efficiency of women as farm managers: profit function analysis in Coˆte d'Ivoire. *Agricultural Economics* 16 (1): 47–53.

Aguilar, A., E. Carranza, M. Goldstein, T. Kilic, and G. Oseni. 2015. Decomposition of gender differentials in agricultural productivity in Ethiopia. *Agricultural Economics* 46 (2015): 311-334.

Alozie, N.O., and P. Akpan-Obong. 2017. The digital gender divide: confronting obstacles to women's development in Africa. *Development Policy Review* 35: 137–160.

Burke, W.J., E. Frossard, S. Kabwe, and T.S. Jayne. 2016. Understanding fertilizer effectiveness and adoption on maize in Zambia. International Development Working Paper 147, Michigan State University Department of Agricultural, Food and Resource Economics, East Lansing.

Burke, W.J., T.S. Jayne, and J.R. Black. 2017. Factors explaining the low and variable profitability of fertilizer application to maize in Zambia. *Agricultural Economics* 48 (1): 115-126.

de Brauw, A., Q. Li, C. Liu, S. Rozelle, and L. Zhang. 2008. Feminization of agriculture in China? Myths surrounding women's participation in farming. *The China Quarterly* 194:327–48.

Falola, T., and N. A. Amponsah. 2012. *Women's roles in sub-Saharan Africa*. Greenwood Publishing: Santa Barbara, CA.

Fuentes, D.O., and H. Wiig. 2009. Closing the gender land gap: the effects of land-titling for women in Peru. Norwegian Institute for Urban and Regional Research Working Paper

No. 2009:120. Oslo.

Gilbert, R.A., W.D. Sakala, and T.D. Benson. 2002. Gender analysis of a nationwide cropping system trial survey in Malawi. *African Studies Quarterly* 6: 223–243.

Goldstein, M., and C. Udry. 1999. Agricultural innovation and resource management in Ghana. Final report to IFPRI under MP17. International Food Policy Research Institute, Washington, DC.

Honig, L., and B.P. Mulenga. 2015. The status of customary land and the future of smallholder farmers under the current land administration system in Zambia. Indaba Agricultural Policy Research Institute Working Paper no. 101. Lusaka, Zambia.

Horrell, S., and P. Krishnan. 2007. Poverty and productivity in female-headed households in Zimbabwe. *Journal of Development Studies* 43 (8): 1351–1380.

Jayne, T.S., N.M. Mason, W.J. Burke, and J. Ariga. 2018. Taking stock of Africa's second-generation agricultural input subsidy programs. *Food Policy* 75 (2018): 1-14.

Jones, A., H. Breuning-Madsen, M. Brossard, et al. 2013. *Soil atlas of Africa*. European Commission, Publications Office of the European Union. Luxembourg.

Karamba, R.W. and P.C. Winters. 2015. Gender and agricultural productivity: implications of the Farm Input Subsidy Program in Malawi. *Agricultural Economics* 46 (2015): 357-374.

Kilic, T., P. Winters, and C. Carletto. 2015. Gender and agriculture in sub-Saharan Africa: introduction to the special issue. *Agricultural Economics* 46 (2015): 281-284.

Larson, D.F., S. Savastano, S. Murray, and A. Palacios-Lopez. 2015. Are women less productive farmers? How markets and risk affect fertilizer use, productivity, and measured gender effects in Uganda. Policy Research Working Paper 7241, Development Research Group. The World Bank, Washington D.C.

Marenya, P.P., and C.B. Barrett. 2009a. Soil quality and fertilizer use rates among smallholder farmers in Western Kenya. *Agricultural Economics* 40.5: 561–572.

Marenya, P.P., and C.B. Barrett. 2009b. State-conditional fertilizer yield response on Western Kenyan farms. *American Journal of Agricultural Economics* 91.4: 991–1006.

Mason, R., P. Ndlovu, J.R. Parkins, and M.K. Luckert. 2015. Determinants of food security in Tanzania: gendered dimensions of household headship and control of resources. *Agriculture and Human Values*, 32(3): 539-549.

Matsumoto, T., and T. Yamano. 2013. Optimal fertilizer use on maize production in East Africa" in *Emerging development of agriculture in East Africa: markets, soil, and innovations*, eds T. Yamano, K. Otsuka, and F. Place. New York, NY: Springer.

Oseni, G., P. Corral, M. Goldstein, and P. Winters. 2015. Explaining gender differentials in agricultural production in Nigeria. *Agricultural Economics* 46 (2015): 285-310.

Sitko, N.J., J. Chamberlin, and M. Hichaambwa. 2014. Does smallholder land titling facilitate agricultural growth? An analysis of the determinants and effects of smallholder land titling in Zambia. *World Development* 64: 791–802.

Smith, W. L., and R. Naylor. 2014. Land institutions and food security in sub-Saharan Africa in *The evolving sphere of food security*, ed R.L. Naylor, 202–35. Oxford, UK: Oxford University Press.

Tiruneh, A., T. Tesfaye, W. Mwangi, and H. Verkuijl. 2001. Gender differentials in agricultural production and decision-making among smallholders in Ada, Lume and Gimbichu Woredas of the Central Highlands of Ethiopia*.* CIMMYT and the Ethiopian Research Organization. El Baton, Mexico.

Udry, C. 1996. Gender , agricultural production , and the theory of the household. *Journal of Political Economy* 104 (5): 1010–1046.

Wooldridge, J. 2002. *Econometric analysis of cross section and panel data*. MIT Press.

Zambia National Farmers' Union (ZNFU). 2013. Time to give women farmers prominence is now. http://www.znfu.org.zm/article/time-give-women-farmers-prominence-now. Accessed June 7, 2018.

Zambia Agriculutral Research Institute (ZARI). 2002. Maize production guide in *Permian Exploration, Boundaries, and Stratigraphy*, eds G.C. Kaitisha, P. Gondwe, M.V. Mukwavi, and G.M. Kaula. Mt. Makulu, Zambia: Soils and Crops Research, Ministry of Agriculture and Cooperatives.

Table 1 – Mean values of yield and determinants by gender group

Variables	Full Sample	Gender group: Female spouses of male head	Female household head	Male
Maize yield (kg/ha)	2268	2811	1723	2413
Farm management				
Total fertilizer application rate(kg/ha)	209	288	151	224
Seed rate	33	41	30	34
Improved (hybrid/OPV) use	0.61	0.57	0.50	0.64
Planted before rain	0.25	0.23	0.24	0.25
Used agroforestry trees	0.03	0.01	0.05	0.02
Weeded twice	0.49	0.40	0.53	0.48
Weeded three times	0.06	0.01	0.06	0.07
Mixed crops	0.04	0.11	0.03	0.04
Province (Share of obs)				
Copperbelt	0.06	0.04	0.06	0.06
Eastern	0.20	0.14	0.22	0.20
Luapula	0.09	0.19	0.05	0.09
Lusaka	0.03	0.07	0.04	0.03
Muchinga	0.08	0.09	0.07	0.09
Northern	0.10	0.24	0.06	0.11
Northwestern	0.05	0.00	0.04	0.05
Southern	0.13	0.03	0.18	0.13
Western	0.10	0.15	0.15	0.09
Education[a] & extension (share of obs)				
Primary school	0.57	0.52	0.63	0.56
Secondary school	0.27	0.29	0.09	0.33
Post-secondary or more	0.05	0.07	0.04	0.05
Received extension advice on conservation farming from Ministry	0.22	0.25	0.16	0.24
Soil characteristics				
Clay <5%	0.18	0.16	0.21	0.17
Clay 15-30%	0.25	0.47	0.24	0.25
Clay 30-45%	0.49	0.29	0.46	0.52
Clay >45%	0.07	0.08	0.09	0.06
pH	5.37	5.24	5.35	5.38
Soil organic matter	1.80	1.81	1.79	1.81
Sample size	1,573	62	1,244	267

Sources: IAPRI – Rural Agricultural Livelihoods and Largest Maize Field surveys (2012). =binary variable where 1=yes. a-"Education" refers to the highest level of education completed by the household head. The manager-specific education for the 62 female spouses of male heads is not explicitly included in our data.

Table 2: Distribution of yields and explanatory variables used in analysis

Variables	Mean	Min	Max	Std. dev.
Maize yield (kg/ha)	2268	2811	1723	2413
Farm management				
Total fertilizer application rate(kg/ha)	209	0	1922	265
Seed rate	33	1	250	30
Improved (hybrid/OPV) use	0.61	0	1	0.49
Planted before rain	0.25	0	1	0.43
Used agroforestry trees	0.03	0	1	0.16
Weeded twice	0.49	0	1	0.5
Weeded three times	0.06	0	1	0.25
Mixed crops	0.04	0	1	0.2
Province (Share of obs)				
Copperbelt	0.06	0	1	0.24
Eastern	0.20	0	1	0.4
Luapula	0.09	0	1	0.28
Lusaka	0.03	0	1	0.18
Muchinga	0.08	0	1	0.28
Northern	0.10	0	1	0.31
Northwestern	0.05	0	1	0.22
Southern	0.13	0	1	0.34
Western	0.10	0	1	0.31
Education[a] & extension (share of obs)				
Primary school	0.57	0	1	0.49
Secondary school	0.27	0	1	0.44
Post-secondary or more	0.05	0	1	0.21
Received extension advice on conservation farming from Ministry	0.22	0	1	0.42
Soil characteristics				
Clay <5%	0.18	0	1	0.39
Clay 15-30%	0.25	0	1	0.44
Clay 30-45%	0.49	0	1	0.5
Clay >45%	0.07	0	1	0.26
pH	5.37	3.1	7.4	0.62
Soil organic matter	1.80	0.09	7.88	0.69
Sample size	1,573			

Sources: IAPRI – Rural Agricultural Livelihoods and Largest Maize Field surveys (2012). =binary variable where 1=yes. a-"Education" refers to the highest level of education completed by the household head. The manager-specific education for the 62 female spouses of male heads is not explicitly included in our data.

Table 3a: Estimated "female effect" on yields from various model specifications

Yield (kg/ha) *Determinants*	(i)	(ii)	(iii)	(iv)	(v)
Female	-555***	-463***	-220**	-227**	-190*
Provincial FE	-	*Yes****	*Yes*	*Yes**	*Yes**
Fertilizer rate	-	-	4.2***	4.2***	4.1***
Seed rate	-	-	19.8***	19.9***	20.1***
Hybrid/OPV seed	-	-	135	144	109
Other Mgmt	-	-	-	*Yes***	*Yes***
Education	-	-	-	-	*Yes*
Soil Organic Matter	-	-	-	-	-
Soil pH	-	-	-	-	-
Clay content					
15-30%	-	-	-	-	-
30-45%	-	-	-	-	-
>45%	-	-	-	-	-
Constant	2,413***	2,682***	722***	722**	607**
N	1573	1573	1573	1573	1573
Chow statistic	-	-	-	-	-
R-squared	0.01	0.06	0.46	0.46	0.46

Sources: IAPRI – Rural Agricultural Livelihoods and Largest Maize Field surveys (2012). *, **, *** Indicates statistical significance at the 10%, 5%, and 1% levels respectively according to standard error estimates robust to arbitrary heteroskedasticity. "Other management" includes early planting, use of agroforestry, number of weedings and other crop mixes. "Education" is three binary indicators for primary, secondary and post-secondary (college or technical school) education and whether they have received extension advice.

Table 3b: Estimated "female effect" on yields from various model specifications

Yield (kg/ha) *Determinants*	(vi)		(vii)		(viii)		(ix)	
Female	-216	-183	-188	-202*	-174	-138	-173	-161
Provincial FE	*Yes****	*Yes*	*Yes****	*Yes*	*Yes****	*Yes*	*Yes***	*Yes*
Fertilizer rate	1.9***	4.7***	2.1***	4.7***	3.7***	4.9***	3.7***	4.9***
Seed rate	18.2**	20.8***	19.2**	20.5***	26.2***	11.3***	25.7***	11.5***
Hybrid/OPV seed	383	-3	377	-11	242	-85	219	-86
Other Mgmt	*Yes*	*Yes*	*Yes*	*Yes*	*Yes****	*Yes*	*Yes****	*Yes*
Education	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Soil Organic Matter	Low (<1.3%)	High (>1.3%)	Low (<1.3%)	High (>1.3%)	-	-	-	-
Soil pH					Low (<5.6)	High (>5.6)	Low (<5.6)	High (>5.6)
Clay content								
15-30%	-	-	114	265*	-	-	192	176
30-45%	-	-	652*	266**	-	-	377**	346
>45%	-	-	548	527**	-	-	675**	445
Constant	944	575**	599	383	461	878**	264	592
N	335	1238	335	1238	1075	498	1075	498
Chow statistic	3.04***		2.59***		2.64***		2.25***	
R-squared	0.34	0.51	0.35	0.52	0.48	0.49	0.48	0.49

Sources: IAPRI – Rural Agricultural Livelihoods and Largest Maize Field surveys, 2012.
*, **, *** Indicates statistical significance at the 10%, 5%, and 1% levels respectively according to standard error estimates robust to arbitrary heteroskedasticity.
"Other management" includes early planting, use of agroforestry, number of weedings and other crop mixes. "Education" is three binary indicators for primary,
secondary and post-secondary (college or technical school) education and whether they have received extension advice. The Chow statistic tests the null hypothesis that there is no structural difference in yield equations on separate soil regimes; a significant statistic rejects this hypothesis in favor of the multiple-regime model.

Table 4: Gender-based differences in pH by province in Zambia

Province	Mean pH on male-managed fields	Difference between pH on male-managed fields and fields managed by: Female heads of household	Female spouses of male heads of household	N
Central	5.42***	-0.11	-0.23***	160
	(0.052)	(0.171)	(0.068)	
Copperbelt	5.13***	0.09	-0.46***	129
	(0.061)	(0.167)	(0.164)	
Eastern	5.66***	-0.02	-0.04	385
	(0.037)	(0.071)	(0.183)	
Luapula	5.07***	-0.05	0.09	134
	(0.076)	(0.166)	(0.262)	
Lusaka	5.67***	-0.03	0.80***	70
	(0.087)	(0.157)	(0.303)	
Muchinga	5.40***	0.17	-0.53***	127
	(0.080)	(0.199)	(0.204)	
Northern	5.07***	-0.03	0.13	163
	(0.043)	(0.094)	(0.180)	
Northwestern	5.44***	0.23	-0.19*	74
	(0.111)	(0.239)	(0.117)	
Southern	5.55***	-0.19	-0.31	194
	(0.064)	(0.165)	(0.262)	
Western	5.11***	-0.12	-0.09	137
	(0.095)	(0.176)	(0.117)	

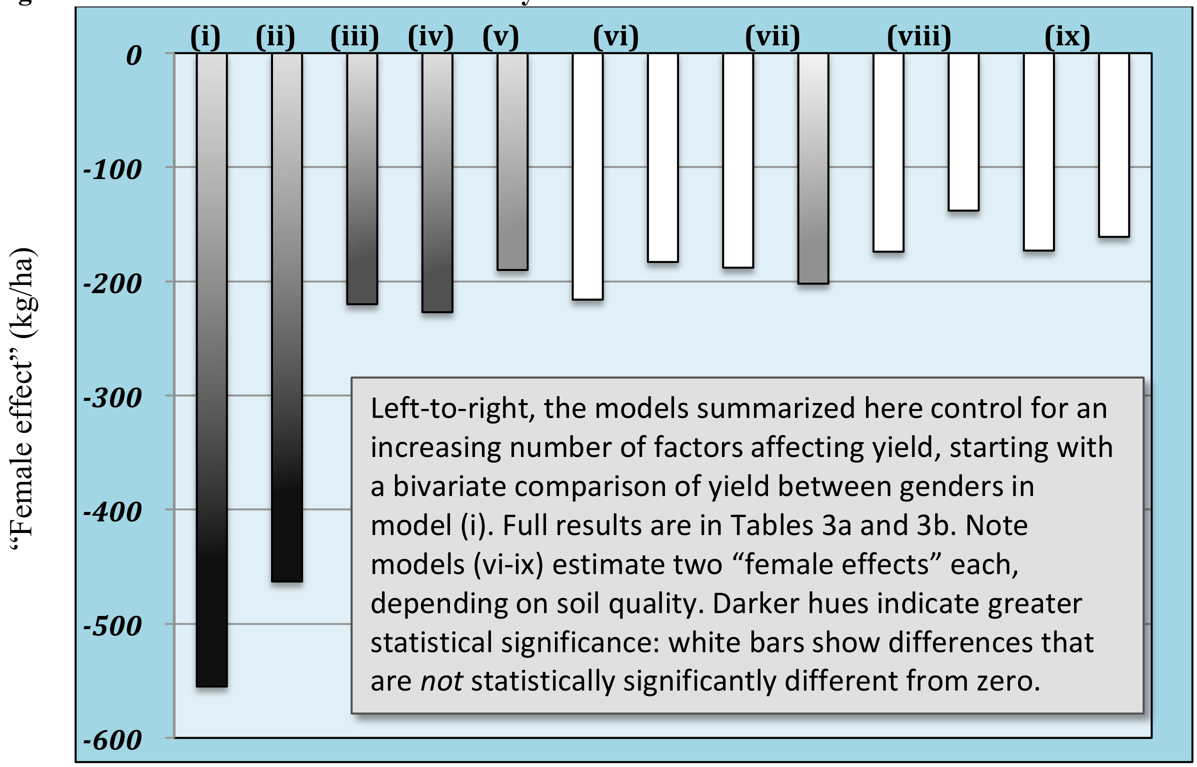
Results estimated separately for each province by regressing pH as the dependent variable on two binary indictors for whether the field is managed by a female head of household or a female spouse of a male head of household. Sources: IAPRI – Rural Agricultural Livelihoods and Largest Maize Field surveys (2012). *, **, *** Indicates statistical significance at the 10%, 5%, and 1% levels respectively

Table 5: Gender-based differences in soil organic matter (SOM) by province in Zambia

Province	Mean SOM on male-managed fields	Difference between SOM on male-managed fields and fields managed by: Female heads of household	Female spouses of male heads of household	N
Central	1.91***	0.11	0.59***	160
	(0.070)	(0.191)	(0.115)	
Copperbelt	2.04***	0.11	-0.07	129
	(0.087)	(0.187)	(0.383)	
Eastern	2.05***	-0.08	0.13	385
	(0.056)	(0.108)	(0.300)	
Luapula	1.80***	0.30**	-0.16	134
	(0.066)	(0.169)	(0.178)	
Lusaka	1.63***	0.25	-0.73*	70
	(0.132)	(0.253)	(0.430)	
Muchinga	1.42***	-0.13	-0.29	127
	(0.076)	(0.155)	(0.186)	
Northern	1.93***	-0.10	0.28	163
	(0.061)	(0.171)	(0.284)	
Northwestern	2.38***	0.01	-0.97***	74
	(0.106)	(0.251)	(0.202)	
Southern	1.46***	-0.04	-0.34*	194
	(0.053)	(0.129)	(0.197)	
Western	1.37***	0.17	0.20	137
	(0.068)	(0.124)	(0.145)	

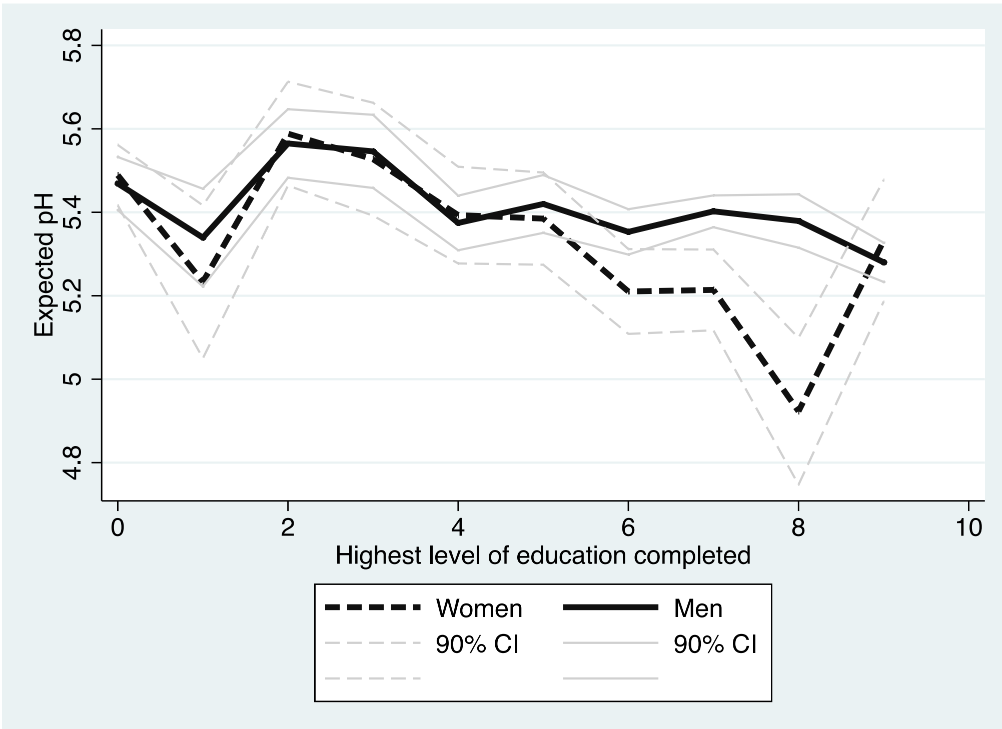
Results estimated separately for each province by regressing SOM as the dependent variable on two binary indictors for whether the field is managed by a female head of household or a female spouse of a male head of household. Sources: IAPRI – Rural Agricultural Livelihoods and Largest Maize Field surveys (2012). *, **, *** Indicates statistical significance at the 10%, 5%, and 1% levels respectively

Fig 1. "Female Effect" estimates on maize yields in Zambia from various models



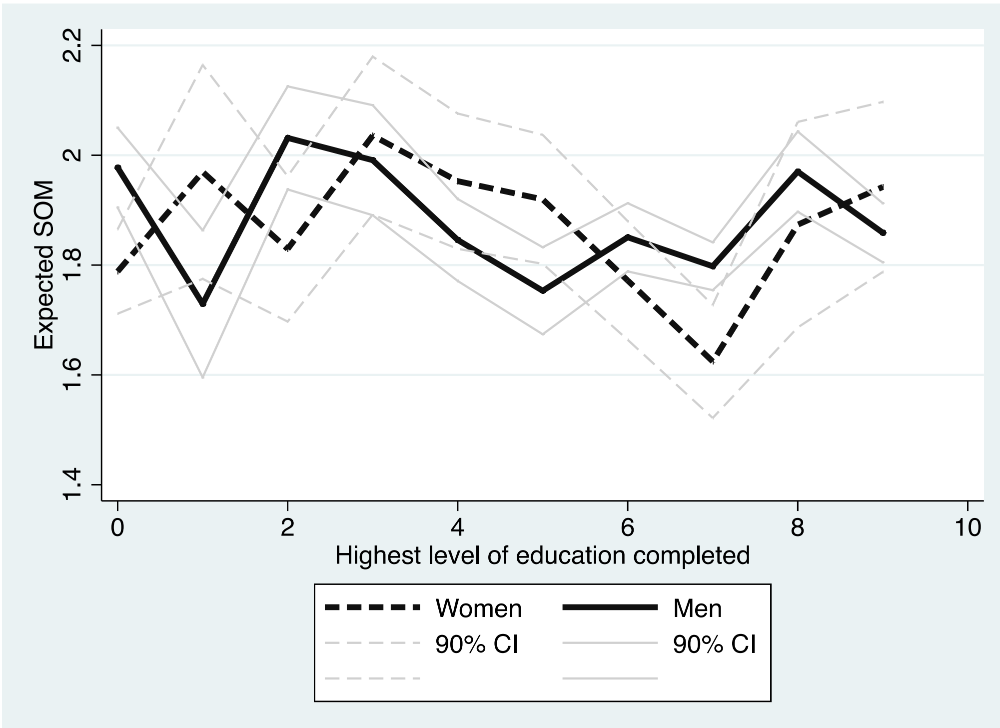
Sources: IAPRI RALS and LMF surveys (2012)

Fig. 2: Expected soil pH by gender of the field manager and education levels among Zambian maize farmers



Sources: IAPRI RALS and LMF surveys (2012)

Fig 3. Expected soil organic matter by gender of the field manager and education levels among Zambian maize farmers



Sources: IAPRI RALS and LMF surveys (2012)

Appendix

Threshold models are estimated such that determinants affect yield in a linear fashion, but these relationships can differ depending on which soil regime crops are planted. For example, if the soil variable is pH, a stylized model would be:

$$yield_i = \begin{cases} \alpha^{(1)} + \beta_1^{(1)} female + \boldsymbol{X\gamma^{(1)}} + u_i \text{ if } pH \leq \theta \\ \alpha^{(2)} + \beta_1^{(2)} female + \boldsymbol{X\gamma^{(2)}} + u_i \text{ if } pH > \theta \end{cases} \tag{A1}$$

The variables *yield, female,* and *pH* and the parameters α and β are defined as in the main text; the $\boldsymbol{\gamma}$ parameters are coefficients on additional determinants ***X.*** The threshold value, θ, can be estimated as that which best fits the data (minimizes squared residuals). Equation A1 is consistent with agronomic literature on the threshold-like relationship between soil characteristics and productivity (Jones et al. 2013, e.g.), and is rooted in theories dating back to the work Carl Sprengel and Justice von Liebig (Sprengel 1838; von Liebig 1840; van der Ploeg, Böhm, and Kirkham 1999). For this study we impose the thresholds found using the same data in Burke et al. (2016), and test for whether these values represent a valid sample splitting in the current context using the Chow statistic to test the null hypothesis there is no difference between regimes – A statistically significant rejection of this hypothesis supports the split-sample, or multiple soil-regime models.

Additional references for the appendix:

Sprengel, C. 1838. *Die lehre von den urbarmachungen und grundverbesserungen (The science of cultivation and soil amelioration)*. Immanuel Müller Publ. Co., Leipzig, Germany.

van der Ploeg, R.R., W. Böhm and M.B. Kirkham. 1999. On the origin of the theory of mineral nutrition of plants and the law of the minimum. *Soil Science Society of America Journal.*

von Liebig, J.. 1840. *Die organische chemie in ihrer Anwendung auf agricultur und physiologie (Organic chemistry in its applications to agriculture and physiology)*. Friedrich Vieweg und Sohn Publ. Co., Braunschweig, Germany.