ABSTRACT

Title of Dissertation: ESSAYS ON DECOMPOSITION ANALYSIS

Amparo Palacios López,
Doctor of Philosophy, 2013

Dissertation directed by: Professor Ramón López
Department of Agriculture and Resource Economics

Chapter 1 shows the implications of credit and labor market imperfections on gender differences in agricultural labor productivity, especially highlighting how both imperfections negatively affect female productivity by discouraging off-farm income generating activities and restricting access to inputs. The paper theoretically models the relationship between gender differences in agricultural labor productivity and market imperfections and it provides empirical evidence consistent with our theoretical model by decomposing the contribution of different factors to such gender differences. We find that agricultural labor productivity is on average 44 percent lower on plots belonging to female-headed households than on those belonging to male-headed households; and that 34 percent of the agricultural labor productivity gap is explained by spillovers from labor market gender differences and 30 percent is explained by gender differences in the use of purchased inputs.

Chapter 2 provides a decomposition analysis of the observed reductions in sulfur dioxide, nitrogen dioxide and ozone concentrations, in the twelve richest European

countries. It quantifies the proportion of the reductions that can be attributed to fiscal policies, trade, and energy taxes. We find that increasing the share of fiscal spending in GDP and shifting the emphasis towards spending in public goods and against non-social subsidies significantly lower the concentrations of sulfur dioxide and ozone but not nitrogen dioxide. At the same time, energy taxes reduce nitrogen dioxide concentrations but have no effect on ozone and sulfur dioxide. Finally trade openness has a direct effect on sulfur dioxide but no effect on nitrogen dioxide or ozone. Our estimates account for time-varying unobserved heterogeneity.

Chapter 3 is the first paper that uses the nationally representative Malawi 2009/2010 dataset. Its purpose is the initial statistical verification of the obtained data and provides a first assessment of agricultural productivity and gender in Malawi. We find that while female-managed plots are, on average, 25 percent less productive, 82 percent of this mean differential is explained by differences in inputs, assets and household characteristics, mainly due to high-value crop cultivation and household adult male labor inputs.

ESSAYS ON DECOMPOSITION ANALYSIS

By

Amparo Palacios López

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in fulfillment
of the requirements for the degree of
Doctor of Philosophy,
2013

Advisory Committee:

Professor Ramón López, Chair
Professor Vivian Hoffmann
Professor Howard Leathers
Professor Kenneth Leonard
Professor Patricio Korzeniewicz, Dean's Representative

Dedication

To my parents Edgar and Amparo

To my daughters Sarah and Anna

To my husband Michael

Acknowledgements

I am grateful to my advisor, Ramón López, for his constant guidance during these seven years, for encouraging me to always be critical of my own work and supporting me in reaching my potential. Thank you for reminding me so many times that what matters is not a paper, but the impact it has in improving this world and for reminding me why I love what I do.

I would also like to thank my committee members – Kenneth Leonard, Howard Leathers, Vivian Hoffmann, and Patricio Korzieniwiecz – for valuable comments to my research.

My time in AREC has been enriched by the friendships that have provided endless support to me during this process. Adan, Asif and Kabir have been my constant companions in this journey and I thank them for their wisdom, support and for all the good moments we spent together. Special thanks to Asif for all the afternoons that he sacrificed helping me with this dissertation, and also for the movies.

My family has been my main support. Thank you to my mom who passed away during my time in AREC for always helping me see things in perspective, for always believing in me and for teaching me to never give up on my dreams. My dad has taught me with his example that with determination it is possible to reach the goals. He has also taught that achieving a degree does not mean anything if it is not put to the service of others. My dad has been the person that has been with me in the hardest moments, he is the one who inspired me to pursue a PhD and continues to challenge me every day. My siblings Edgar and Xochitl have always been my best friends and have given me their love and support, thanks for being always there to listen to me and for being patient with my absences from family activities.

I have been able to work on my dissertation because I have had a network of friends and family that have supported and helped me and my family throughout this journey. I would like to acknowledge some of them although the list is long: my godparents Jerry and Carolyn Parr, Carol Blythe and Rick Goodman, my aunts Miriam and Yolanda, my cousins Karla and Morena, my parents in-law Eleonora and Willi Settels, my sisters in law Uta, Katrin and Angelika and Maria and Peter Donges, and the friends from Calvary Baptist Church and Festival Church. All of you have encouraged me and supported me in so many ways that I am eternally indebted to you.

My colleagues at the LSMS-team of the World Bank have supported this dissertation by allowing me to use the data before its dissemination. Special thanks to Talip Kilic for being my mentor and for introduced me the gender and agriculture field.

Thanks to my husband Michael, I began this journey because you supported my dream, I stayed in graduate school because you gave me the strength to continue even when I felt that I did not have energy, I am finishing because you have been here with me. This dissertation is yours as much as it is mine.

Finally thanks to my two beautiful daughters Sarah and Anna Xochitl, you are my inspiration, you have been an oasis in the middle of the desert and I am looking forward to continue watching you grow and being your guide in this world.

Table of Contents

List of Tables

List of Figures

Market Imperfections and Gender Differences in Agricultural Productivity in Malawi

By: Amparo Palacios-López

1 Introduction

Two key features are prevalent in the Agricultural sector in Sub-Saharan Africa. The first one is the dominance of female labor in agriculture and the second one is the existence of a gender gap in agricultural productivity. The feminization of agriculture is evident in Sub-Saharan Africa, where women make up a higher proportion of agricultural labor than males, potentially ranging from 30 to 80 percent (UNECA 1982, FAO 1984, Doss 2011, Doss et al. 2011). Nevertheless, labor conditions in the rural sector are disadvantageous for women. The presence of gender differences in the rural sector is not surprising in the context of Malawi. In a 2004 survey, rural wages were 35% lower for females than males. Furthermore, around 89% of employed women are engaged in part time off-farm activities, in contrast to 67% for male. Women participating in rural wage employment tended to be concentrated in lower skill activities- about 61.4 % - in contrast to the corresponding figure of 37% for males.

The gender differentials in agricultural productivity range from 4 to 40 percent conditional on the country, the representativeness of the data, the type of crop, and the composition of households among other variables (Akresh, 2005; Alene et al., 2008;

Gilbert et al., 2002; Goldstein and Udry, 2008; Moock 1976; Peterman et al., 2011; Oladeebo and Fajuyigbe, 2007; Quisumbing et al., 2001; Saito et al., 1994; Tiruneh et al., 2001; Udry, 1996; Vargas Hill and Vigneri, 2011). The disparity of large participation of women in agriculture, and the significant gender gap in agricultural productivity are the prime motivations for this study. The implications of both characteristics are that the agricultural sector may lag behind its potential in terms of productivity, and that a gender dimension exists within the issue at hand. Thus, it is important to understand the underlying reasons behind the differences in agricultural productivity between female and male farmers.

Several key reasons for the observed gender gap in agricultural productivity have been identified in the literature: gender differences in (i) access and use of agricultural inputs, (ii) tenure security and related investments in land and improved technologies, (iii) market and credit access, (iv) human and physical capital, and (v) informal and institutional constraints affecting farm/plot management and marketing of agricultural produce (Peterman et al. 2011). Cultural roles that are assigned to males and females regarding domestic duties as well as other factors that may underlie the gender segregation in crop production (i.e. staple vs. cash crop cultivation, high-yielding vs. low-yielding variety cultivation, etc.) could be thought of as informal institutional constraints. However, the relationship between labor market discrimination and the observed gender gap in agricultural productivity has received little attention in the literature.

This study focuses on gender differences in agricultural labor productivity. This is an important deviation from the prevalent literature that concentrates on agricultural land productivity. The principal asset of the poor is their labor power, and a disproportionate number of poor households are headed by women[1]. Direct increases in labor productivity as well as labor opportunities raise the income-earning capacity which is particularly important for female-headed households.

The literature has acknowledged the presence of multiple market failures in agriculture especially in the labor market. Labor supply behavior is affected by risk, search and transaction costs, locational preferences, gender preferences, and gender discrimination (Barret, 1996; Biswanger and Rosenzweig, 1986). This drives a wedge between the marginal product of labor and the prevailing market wage rate for the same type of labor (Barret et al. 2008). The labor allocations resulting from this deviation from the equilibrium condition may be welfare-maximizing, thus individuals may still be optimizing their allocation of labor. As an important contribution of this paper we show that the wedge between the marginal product of labor and the prevailing market wage rate varies by gender and is generally larger for women. We argue that this may be explained by market imperfections.

The literature on gender discrimination has typically followed one of two avenues. One explores labor market discrimination in terms of wages in off-farm activities (O'Neill & O'Neill, 2006, Fortin, 2006). The other explores gender differences in the agricultural sector with a focus on issues related to inputs, credit access, market access, and cultural constraints (Peterman et al. 2010). Our study contributes to the

---

[1] In 2004/05 nationally representative household survey of Malawi, 78 percent of rural households were poor, 25 percent of these poor households were female-headed.

literature by both exploring the effects of gender discrimination in the labor market on agricultural productivity, and by analyzing the implications for the agricultural productivity gender gap.

This is one of the first studies to both theoretically and empirically decompose the sources of gender differences in *agricultural labor productivity* and relate them to credit and labor market imperfections. There are several key implications. Credit markets may treat women and men differently in a discriminatory fashion which causes women to have less access to purchased inputs (FAO 2011b). Labor market imperfections result in women receiving lower wages from off-farm activities than men (FAO 2011b, Hertz et al. 2009, Doss 2011, Doss et al. 2011). This spills over to agricultural labor productivity causing women to work less in off-farm activities and more on the farm than men therefore creating productivity differentials. Furthermore the additional off-farm time burden due to engagement in off-farm work is higher for women than men due to the burden of domestic activities women face which lead to a greater allocation of female labor towards on-farm work causing their agricultural labor productivity to be lower than men's. In addition, differences in the household's asset endowment and human capital may be the result of the long term effects of such market imperfections.

The theoretical model of household farm productivity in this study deviates from the literature in a couple of ways. It examines agricultural labor productivity subject to the head-of-household's gender and traces the effects of labor market imperfections and credit market constraints on productivity. The model's inclusion of gender specific off-farm time burden due to engagement in off-farm activities borrows from studies by López (1984, 1986) who examines how time allocation between on-farm and off-farm work has different connotations on utility as a consequence of commuting time required by off-farm work. He shows how the consideration of commuting time leads to a model in which the household acts as having different preferences between on-farm and off-farm work, even if preferences are defined purely in terms of leisure. In this setting the optimization problem of the household becomes non-separable and thus the production decisions would be directly linked to the household's characteristics and consumption decisions. Our theoretical set up models household preferences as additive and increasing in the present value of earnings and leisure (Eswaran and Kotwal, 1986). Our model predicts that the gender specific off-farm time burden due to engagement in off-farm activities leads to labor productivity differences between men and women, with women as heads of household devoting a higher proportion of their time to agricultural activities.

Predictions derived from the model are empirically illustrated using nationally representative data from Malawi. We explore gender differences in productivity at the plot level. We specifically compare female managed plots to male managed plots, noting that the respective managers are also heads of their households. 23 percent of the plots in the sample are managed by female heads of household. It is important to note that female heads of household include mostly divorced, separated, single, or never married women. Altogether this represents 86 percent of the female heads of household in the sample. The remaining 13 percent of female heads of household are married. Thus our examination of gender differences mostly pertains to female headed households in which a male spouse is absent. Therefore the ensuing results may not necessarily apply to the case of female plot managers which live with their spouses.

The implemented econometric approach adapts a decomposition method from labor economics, most notably in the analyses of the gender wage gap, union wage gap, and growing wage inequality (Oaxaca, 1973 and Blinder, 1973). Our study specifically decomposes the average differences in agricultural labor productivity between plots belonging to male-headed households and those belonging to female-headed households into four effects: (i) labor market effect, (ii) purchased inputs effect, (iii) endowment effect, and (iv) pure marginal productivity effect which is the gender differences in coefficients of the various factors of production and household characteristics.

The labor market effect refers to the portion of the agricultural labor productivity gender gap that is driven by gender differences in the number of hours the manager works on the plot. This effect may in part reflect the direct impact of labor market imperfections that affect men's and women's allocation of labor between on-farm and off-farm work differently, with men allocating more time to off-farm work than women. The purchased inputs effect refers to the portion of the gap explained by gender differences in the use of purchased inputs, such as fertilizer, pesticides, hired labor and agricultural implements; this may reflect the direct impact of credit constraints on agricultural productivity, which affects men and women differently, limiting further the capacity of women to buy agricultural inputs in comparison to men. The purchased inputs effect may also capture the indirect effect of the labor market imperfections, by capturing the degree in which the credit constraint is relieved through off-farm income. The endowment effect includes gender differences in household characteristics and assets owned by the household; this effect may in part reflect the long term impact of labor market imperfections and credit market constraints that limit the capacity of accumulation of human and physical capital and that may affect women more than men. If in fact the endowment effect is proven to be significant, this would be a manifestation of the non-separability nature of a household's decisions, thus production decisions become directly linked to the household's characteristics and consumption decisions. Finally, the pure marginal productivity effect is the portion of the agricultural labor productivity gender gap that is driven by gender differences in the set of coefficients of all the covariates included in the regressions, this effect may also encompass the long term effects of labor and credit market constraints on agricultural production as well as cultural and institutional constraints that may underlie gender segregation in crop production.

Four key results can be derived from our study. First, agricultural labor productivity is on average 44 percent lower on plots belonging to female-headed households (*female plots*) than on those belonging to male-headed households (*male plots)*. Second, 34 percent of the agricultural labor productivity gap is explained by spillovers from labor market gender differences and 30 percent is explained by gender differences in the use of purchased inputs. Third, on average, agricultural land productivity on female plots in Malawi is 25 percent lower than on male plots. Fourth, when analyzing the gender gap at different deciles of the agricultural labor productivity, it can be observed that the gender gap increases across the distribution, reaching a maximum value of 54 percent at the 70^{th} percentile of the labor productivity distribution.

The rest of the paper is structured as follows. Section 2 presents a theoretical model that aims to explain the gender differences in agricultural productivity in developing countries. Section 3 explains the strategy used to empirically verify the model's predictions and presents the mean decomposition methodology used. Section 4 describes the data and Section 5 shows the decomposition results. In section 6 we present the sensitivity analysis and Section 7 offers concluding remarks and expands on the policy implications of our findings.

2 Theoretical Model

Farmers in developing countries face budget and working capital constraints. Given that they experience limited access to credit, they can alleviate credit constraints by generating income from off-farm labor activities. The labor allocation across on-farm and off-farm work, as well as the type and quantity of non-labor agricultural inputs chosen, are central to the productivity of the farm. We develop a model that characterizes the decisions made by different households taking into account the constraints faced a priori (choice of labor, on-farm or off-farm), which in turn affect the allocations of resources ex post (labor and non-labor inputs used in the farm and consumption decisions). We assume the presence of credit market failures, as well as the existence of gender specific off-farm time burden and gender based differential treatment in the labor market that determine the allocation of labor across on-farm and off-farm work. That is, given the presence of credit constraints for example, the decisions made by the household will differ depending on its composition and/or on the opportunities available to men and women, implying that choices made by a female-headed household will be different from the choices made by a male-headed household.

2.1 Household Welfare Maximization

We posit that all households have the same preferences. We assume the utility function is additive and increasing in the present value of earnings (*Y)* and leisure *(l)* (Eswaran and Kotwal, 1986).

$$U(Y, l; Z_{hh}) = Y + u(l) \tag{1}$$

where Z_{hh} is a vector of exogenous household characteristics.

The household allocates its time endowment (*H)* between leisure (*l)*, on-farm labor (L_f), and off-farm labor (L_o). We model the existence of a gender specific “off-farm time burden” due to engagement in off-farm activities (López, 1984, 1986) which is determined by the structure of the household. “Off-farm time burden” as defined here includes commuting time as well as actual off-farm work time. The intuition is as follows. An increase in off-farm time is not simply an equivalent subtraction of hours spent on farm work or leisure. There is an additional time cost incurred due to several reasons. For instance households may have to alter their schedule to accommodate off-farm work. This cost of re-organization of activities may increase with every increase in time spent in off-farm activities. Furthermore, there may be synergies between on-farm work and household work including child bearing. Thus a reduction in time allocated to the farm may incur an additional cost increasing with every additional unit of time spent in off-farm work. Importantly such off-farm time burden may vary by gender. Female headed-households may face higher commuting time burden due to household care responsibilities that are culturally assigned to women (child care, cooking, getting water, etc.).

Given the above considerations, and contrary to the standard practice, we may regard the time constraint faced by the households as non-additive as follows:

$$H = l + L_f + g(L_o) \qquad (2)$$

where $g(L_o)$ is “off-farm time burden” and is equal to $\alpha_o + \alpha_1 L_o$.

The gender specific off-farm time burden has two parameters, α_o a fixed time incurred when the household participates in off-farm activities, which is the same for both male and female headed households and α_1 that represents the portion that is determined by the degree of household care activities, α_1 is greater than or equal to 1, with 1 implying that the household has lower household care responsibilities. We assume that the gender specific off-farm time burden due to engagement in off-farm work is higher for female heads of household thus $\alpha_1^F > \alpha_1^M$, with F and M denoting female- and male-headed households respectively.

The production process requires the use of two variable inputs: labor (L_f) and non-labor inputs (X). Non-labor inputs include inputs such as inorganic fertilizer, improved seeds, and traditional seeds. For simplicity we assume that the farmer uses only household labor. The production function is presented as follows: $f(X, L_f; Z_{prod})$ where Z_{prod} is a vector of exogenous farm characteristics.

We define net income as equal to the sum of revenues from all sources minus household expenditures:

$$Y = w*(L_0) + pf(X, L_f; Z_{prod}) - X*r \qquad (3)$$

where *w* is the wage received in off-farm activities, *r* is the price of non-labor input *X*, and *p* is the price of output. All prices are exogenous.

In addition, we assume that the household faces a working capital (liquidity) constraint where the inputs purchased are less than or equal to the amount borrowed plus the income from off-farm activities. The implicit assumption behind this constraint is that all expenses are incurred at the beginning of the production period.

$$r*X \leq B + wL_o \qquad (4)$$

Thus a household maximizes its utility by allocating labor between on-farm, off-farm work, and deciding how much non-labor input they will use by solving the following maximization problem:

$$\max_{L_o, L_f, X} V = w*(L_0) + pf(X, L_f; Z_{prod}) - X*r + u(H - L_f - (\alpha_o + \alpha_1 L_o)) \qquad (5)$$

Subject to

$$r*X \leq B + wL_o$$

From the optimization problem above, we obtain the following Lagrangean:

$$\max_{L_o, L_f, X} V = w*(L_0) + pf(X, L_f; Z_{prod}) - X*r + u(H - L_f - (\alpha_o + \alpha_1 L_o)) + \psi(B + wL_o - X*r) \qquad (5a)$$

where ψ is the shadow price of the working capital constraint.

Monotonicity of the utility function implies that the working capital constraint will be binding.

The First Order Conditions (FOC) are:

$$\frac{\partial V}{\partial L_o} = w - U_l \alpha_1 + \psi w \leq 0 \tag{6a}$$

$$\frac{\partial V}{\partial L_f} = -U_l + pf_{L_f} \leq 0 \tag{6b}$$

$$\frac{\partial V}{\partial X} = pf_x - r - \psi r \leq 0 \tag{6c}$$

$$\frac{\partial V}{\partial \psi} = B + wL_o - rX \leq 0 \tag{6d}$$

where U_l denotes marginal utility of leisure and f_x and f_{L_f} denote the marginal productivity of input X and on-farm labor respectively.

2.2 Implications for Labor Allocation

Under standard competitive capital and labor markets, the gender specific off-farm time burden due to engagement in off-farm work does not affect the labor allocation decisions of the household, and thus the model is separable. Therefore, the household's production decisions are independent from its consumption decisions and its composition. From the FOC, under competitive capital and labor markets, the household would follow the standard allocative efficiency rule to allocate labor between on-farm and off-farm work, and household characteristics will not play any role in the labor supply of the household. Thus, the household will allocate labor according to (from 6a and 6b):

$$w = pf_{L_f} \tag{7}$$

We now assume that the wage paid for off-farm activities is different for men and women, thus $w^F = w(1-\phi)$, where ϕ represents the differential treatment of the labor market towards men and women and takes values between 0 and 1, with 0 representing a non-discriminatory labor market and 1 representing a discriminatory market. We speculate that the difference in wages reflects some type of discrimination which results in fewer opportunities for women than for men in the off-farm labor market.

Additionally, the shadow price of the working capital constraint $\psi(Z_{hh})$ is a function of the household characteristics; hence it will be different for male and female headed households. Given that the working capital constraint might be binding, the existence of gender differences in wages and the gender specific off-farm time burden due to engagement in off-farm work, the household will allocate labor between on-farm and off-farm work as follows (from 6a and 6b):

Male headed households:

$$\frac{w(1+\psi(Z_{hh}^M))}{\alpha_1^M} = pf_{L_f}^M \tag{8}$$

Female headed households:

$$\frac{w(1-\phi)(1+\psi(Z_{hh}^F))}{\alpha_1^F} = pf_{L_f}^F \tag{9}$$

Consequently, the household will allocate labor towards off-farm work until the net benefit from off-farm work is equal to the net benefit from on-farm work. Allocation of labor towards off-farm work will alleviate the working capital constraint, which can be shown through the effect of $(1+\psi)$ on equation (8). The value of every hour worked in off-farm activities not only provides income but also relieves the working capital constraint. The marginal productivity of labor is an endogenous function of household characteristics, preferences, assets and labor market discrimination. In this setting the model becomes non-separable.

Equations (8) and (9) show us that the allocation of labor in female-headed households tends to be towards farm activities despite lower farm labor productivity of women compared to men.

The level of input X used also depends on the degree of the household's credit constraint. From (6c) we know that the more constrained the households are the less they will invest in X:

$$r(1+\psi) = pf_X \qquad (10)$$

The total effect of α_1 on agricultural labor productivity is ambiguous. Farm labor has a positive relationship with α_1. Less labor is allocated to off-farm work as α_1 increases. On the other hand the use of non-labor inputs decreases with α_1 because less off-farm income is received, thus the working capital constraint is exacerbated. The total effect of α_1 on agricultural labor productivity depends on which agricultural input (labor or non-labor) has a higher impact. The discriminatory effect due to differential treatment toward women in the labor market has also a negative effect on off-farm labor. The total effect of α_1 and labor market discrimination will cause female-headed households to allocate more time to on-farm work when compared to male-headed households. To sum up, the key mechanism is that fewer off-farm opportunities for women due to gender specific off-farm time burden as well as labor market discrimination leads to lower off-farm income, essentially exacerbating any pre-existing liquidity constraints faced by female-headed households. The end result is female-managed plots are quite likely to have fewer non-labor inputs, ultimately reducing their productivity.

2.3 The Gender Gap

Dividing (8) over (9) we obtain a measure of differences in labor agricultural productivity (gender gap in agricultural labor productivity):

$$\frac{f_{L_f}^M}{f_{L_f}^F} = \underbrace{\frac{\alpha_1^F}{\alpha_1^M(1-\phi)}}_{\text{labor market effect}} \underbrace{\frac{(1+\psi(Z_{hh}^M))}{(1+\psi(Z_{hh}^F))}}_{\text{liquidity constraint effect}} \qquad (11)$$

where $f_{L_f}^M$ and $f_{L_f}^F$ are the marginal productivity of labor in male- and female-headed households respectively.

The total labor market effect (gender discrimination and the gender specific off-farm time burden) has a positive effect on the gender gap in agricultural labor productivity. The first term in equation (11) is unambiguously greater than 1 given that $\alpha_1^F > \alpha_1^M$, and $0 \leq \phi \leq 1$ which implies that male agricultural labor productivity is higher than that of female-headed households. Female-headed households will still remain in farm activities even when the return is lower than the return in alternative off-farm activities. This is because they have to account for specific household labor activities they are required to fulfill, and for the differential treatment with regards to off-farm labor opportunities.

The second term in equation (11) reflects the effect of liquidity constraints on the gender gap in agricultural labor productivity, which is ambiguous. The difference in response to liquidity constraints may be different for male- and female-headed households, but the direction of the total effect remains an empirical question. Higher liquidity constraints increase the value of working in off-farm activities, thus might encourage the allocation of labor allocation in off-farm work. We can assume that the effect of the liquidity constraint does not compensate the effect of the labor market (if both effects have opposite signs), in which case, male-headed households will have higher productivity than female headed households and the gender gap will be positive.

2.4 Propositions

The following two propositions summarize the main predictions of the theoretical model

(i) Higher gender specific off-farm time burden due to engagement in off-farm work and labor market discrimination in capital constrained households lead to higher availability of on-farm labor and lower availability of off-farm labor and therefore less off-farm labor income. A decrease in off-farm labor income reduces liquidity that might be used to buy non-labor inputs. This reduced liquidity effect may cause greater need for credit which in the context of credit market imperfections imply a larger negative impact of such imperfections. **Labor market restrictions may exacerbate the effect of credit market imperfections, in agricultural households, reducing off-farm labor and income and thus the access to non-labor inputs**.

(ii) Female headed households have fewer opportunities for obtaining off-farm income and thus will have lower access to non-labor inputs and therefore **lower agricultural labor productivity.**

3 Empirical Analysis

The econometric approach we use has been utilized in labor economics as part of the analyses of the gender wage gap, union wage gap, and growing wage inequality (O’Neill & O’Neill, 2006, Fortin, 2006). We use the mean decomposition methodology to look at the differences in agricultural productivity for male- and female-headed households. The proceeding subsections will provide details of the methodology.

Regression-based decomposition methods have been widely utilized in labor economics following the seminal papers of Oaxaca (1973) and Blinder (1973). Extensive use of Oaxaca-Blinder regression-based mean decomposition among applied economists over time extended its application to the decomposition of distributional statistics. It is, however, acknowledged that the questions attempted to be addressed by this method require a strong set of assumptions (Fortin et. al., 2011). We return to these assumptions after describing the decomposition.

Decomposition methods follow a partial equilibrium approach, where observed outcomes for one group can be used to construct various counterfactual scenarios for the other group (Fortin et. al., 2011). Another characteristic is that while decompositions are useful for quantifying, purely in an accounting sense, the contribution of various factors to a difference in an outcome across groups or a change in an outcome for a particular group over time, they are based on correlations, and hence cannot be interpreted as estimates of underlying causal parameters (Fortin et. al., 2011). However, decomposition methods do document the relative quantitative importance of factors in explaining an observed gap, thus suggesting priorities for further analysis and, ultimately, policy interventions (Fortin et. al., 2011).

We regress Y, the log of value of output per hectare (land productivity) for male- (M) and female- (F) headed household plots, on its determinants as expressed by the following equations:

$$Y^{M} = \beta_0^{M} + L^{M}\beta_L^{M} + \sum_k X_k^{M}\beta_k^{M} + \sum_j Z_j^{M}\delta_j^{M} + \varepsilon^{M} \quad (12a)$$

$$Y^{F} = \beta_0^{F} + L^{F}\beta_L^{F} + \sum_k X_k^{F}\beta_k^{F} + \sum_j Z_j^{F}\delta_j^{F} + \varepsilon^{F} \quad (12b)$$

where L is the number of hours of managerial labor per hectare*;* X is a vector of k purchased inputs (pesticides, organic and inorganic fertilizer, hired labor, agricultural implements, improved seeds); Z is a vector of characteristics of the household that includes human and physical capital (wealth, land assets, household composition, location of the household and location of the plot, access to off-farm income and transfers); β_0, β_L, β_k, δ_j are the associated vector of intercept and slope coefficients for male and female headed households; and ε is the error term under the assumption that $E(\varepsilon^M) = E(\varepsilon^F) = 0$. The decomposition of the gender gap in agricultural land productivity is presented in Appendix A.

3.1 Mean Decomposition Of Labor Productivity

We use the resulting vector of coefficients from the land productivity regressions indicated in equations (12a) and (12b) to create a measure of labor productivity (value of output per hour of managerial labor). We create measures of labor productivity in logarithm form by subtracting labor from land productivity $\ln(\frac{Y}{L}) = \ln(Y/ha) - \ln(L/ha)$.

We use equations (12a) and (12b) and subtract L from both sides of the equation.

$$E(Y^M) - E(L^M) = \beta_0^M + E(L^M)\beta_L^M + \sum_k E(X_k^M)\beta_k^M + \sum_j E(Z_j^M)\delta_j^M - E(L^M) \tag{13a}$$

$$E(Y^F) - E(L^F) = \beta_0^F + E(L^F)\beta_L^F + \sum_k E(X_k^F)\beta_k^F + \sum_j E(Z_j^F)\delta_j^F - E(L^F) \tag{13b}$$

The *gender gap* in labor productivity "D_L" is expressed as the mean outcome difference:

$$D_L = [E(Y^M) - E(L^M)] - [E(Y^F) - E(L^F)] \tag{14}$$

Substituting (13a) and (13b) into (14) and adding and subtracting $E(L^M)\beta_L^F$, $\sum_k E(X_k^M)\beta_k^F$ and $\sum_j E(Z_j^M)\delta_j^F$, we decompose the gender gap in labor productivity into the following four components:

$$D = \underbrace{[E(L^M) - E(L^F)](\beta_L^F - 1)}_{\text{labor market effect}} + \underbrace{\sum_k [E(X_k^M) - E(X_k^F)](\beta_k^F)}_{\text{purchased inputs effect}} + \underbrace{\sum_j [E(Z_j^M) - E(Z_j^F)]\delta_j^F}_{\text{household endowment effect}} +$$

$$\underbrace{E(L^M)[\beta_L^M - (\beta_L^F)] + \sum_k E(X_k^M)[\beta_k^M - \beta_k^F] + (\beta_0^M - \beta_0^F) + \sum_j E(Z_j^M)[\delta_j^M - \delta_j^F]}_{\text{pure marginal productivity effect}} \tag{15}$$

In practice, we estimate equation (12a) and (12b) using the value of output per hectare as the outcome variable. We use the resulting vector of coefficients, in combination with the mean values for each covariate of the male and female samples to compute the components of equation (15).

The difference between the decomposition of the gender gap in terms of land productivity (Appendix A) and labor productivity is that the disparity in hours worked by men and women is weighted by $(\beta_L^F - 1)$, which is the measure of the elasticity of labor in female plots. Hence, in this case, the fact that women work more than men exacerbates the average labor productivity gap between men and women. In contrast, the increase in on-farm work by women relative to men actually increases land productivity albeit in an inefficient way.

The first component of equation (15) is the *labor market effect*, i.e. the portion of the gender gap driven by differences in quantities of labor allocated to on-farm work by the head of household. The second component is the *purchased inputs effect*, the portion of the gender gap that is explained by differences in levels of use of inputs that have to be bought such as fertilizer, pesticides, seeds, agricultural implements, and/or hired labor. The third component, the *household endowment effect* is comprised by differences in levels of observable characteristics of the household, including human and physical capital. The fourth component is the *pure marginal productivity effect* and corresponds to the portion of the gender gap explained by differences in the coefficients of each observable covariate included in L and in the *X* and *Z* vectors; as well as differences in the constant between male- and female-headed households.

The graphical representation of the gender gap in land and labor productivity and their respective components is presented in Figure 1. It presents the value of output per hectare ($Y=y/ha$) at each level of managerial labor hours per hectare ($L=l/ha$). Curve $Y^M\binom{\text{Male Endowment}}{\text{Male Purchased Inputs}}$ and curve $Y^F\binom{\text{Female Endowment}}{\text{Female Purchased Inputs}}$ are derived from the male and female regressions respectively. Curve $Y^F\binom{\text{Female Endowment}}{\text{Male Purchased Inputs}}$ is derived using the coefficients from the female regression and the male average value of the variables included in the purchased inputs vector; this curve represents the level of agricultural productivity that a female-headed household would have if it had the same level of purchased inputs as a male-headed household, but the female level of endowment. Similarly, curve $Y^F\binom{\text{Male Endowment}}{\text{Male Purchased Inputs}}$ is derived using the coefficients of the female regressions and the male average level of the variables included in the endowment and purchased inputs vectors. L^M and L^F are the observed average levels of managerial labor for the male and female samples respectively.

Point A is the observed average male land productivity and point D is the observed average female land productivity. Point E is the productivity that women would get if they would work the same number of hours as men. Point C is the land productivity that women would get if they had only the same level of purchased inputs as men, but the female endowment and would work the same number of hours as men, while point B is the land productivity that women would attain if they had access to the same level of endowment and purchased inputs and worked the same number of hours as men.

The observed gender gap in land productivity is equal to the distance between Y^A and Y^D and can be decomposed into the four components presented in equation (A.4): the pure marginal productivity effect is the distance Y^A-Y^B, the endowment effect Y^B-Y^C, the purchased inputs effect Y^C-Y^E and the labor effect Y^E-Y^D. In the context of land productivity, the fact that women work more hours in average than their male counterparts has the effect of reducing the gender gap, counteracting the purchased inputs, endowment and pure marginal productivity effects.

The average labor productivity is obtained by dividing Y over L ($\frac{Y}{L}$) which is equal to the slope of a straight line from the origin to any point on the land productivity curve.

The observed gender gap in labor productivity is equal to $D^L = \frac{Y^A}{L^M} - \frac{Y^D}{L^F}$ which is decomposed into the four components of equation (15): the pure marginal productivity effect is equal to $\frac{Y^A}{L^M} - \frac{Y^B}{L^M}$, the endowment effect is $\frac{Y^B}{L^M} - \frac{Y^C}{L^M}$, the purchased inputs effect $\frac{Y^C}{L^M} - \frac{Y^E}{L^M}$ and the labor effect is $\frac{Y^E}{L^M} - \frac{Y^D}{L^F}$. For comparison purposes, it is important to note that the average productivity in D is equal to the average productivity in D*. The labor effect is positive unlike the labor effect in the decomposition of the land productivity. In the case of labor productivity, the labor effect amplifies the gender gap.

This is consistent with the theoretical model which predicts that the gender gap in labor productivity will be larger than the gender gap in land productivity given that the labor market imperfections spillovers increase the allocation of labor to the farm. As discussed earlier, this increases production in an inefficient way. Figure 1 illustrates this prediction.

We attempt to use equation (15) to decompose the gender gap in agricultural labor productivity and relate it to the components presented in equation (11): (i) the labor market effect may in part reflect the direct effect of labor market imperfections over agricultural production, (ii) the purchased inputs effect may represent the direct effect of liquidity constraints over agricultural production as well as the indirect effect of labor market imperfections, (iii) the endowment effect may reflect the long term effects of credit and labor market imperfections that affect men and women differently and that limit the capacity to accumulate physical and human capital over time, effect that may limit women more than men; and (iv) the pure marginal productivity effect may also represent the long term impact of credit and labor market imperfections as well as gender differences in the way households make decisions, due to institutional and cultural constraints that define the roles of males and females regarding domestic duties and gender segregation in crop production.

3.2 Empirical Issues

The decomposition methods described above are valid only under certain assumptions. Fortin et al. (2011) present a detailed account of the assumptions required to estimate the population parameters of interest. Two crucial assumptions for the validity of aggregate decomposition are (i) overlapping support and (ii) ignorability.

The overlapping support assumption rules out cases where observable and unobservable covariates may be different across the two groups. Hence "overlap" refers to the similarity of the covariate distributions of both subpopulations. It implies that no single value of the covariates (X, Z, L) attain specific values ($X = x, Z=z, L=l$) or $\varepsilon = e$ exists to identify female plot management.

Ignorability refers to the random assignment of female plot management conditional on observable attributes. Specifically we worry that our male and female managed plots may not be randomly assigned. Ignorability allows us to assume that we have enough controls and thus, conditional on these controls, our assignment of female plot management is essentially randomized. It rules out what we typically call "self-selection" based on unobservables. The additional essential assumptions required by detailed decomposition of the individual contribution of each covariate include additive linearity and zero conditional mean. The latter implies that ε is independent of the explanatory variables. In other words, we assume that there is no unobservable heterogeneity that jointly determines the outcome and observable attributes. The former assumes a linear functional form.

In exploring the existence and extent of the gender gap in a multivariate framework, the validity of findings largely depends on the plausibility of the ignorability and zero conditional mean assumptions, i.e. the extent to which the estimation strategy addresses possible unobservable household-/plot-level heterogeneity that jointly determines plot agricultural productivity and observable covariates, including whether a plot belongs to a household headed by a female. While the most rigorous studies on the gender gap recognize the need for an instrumental variable strategy to deal with potentially endogenous observables, recovering instrumental variables that predict endogenous covariates without directly influencing the outcome is often not possible. We attempt to lend as much support to the assumptions of overlapping support, ignorability, and zero conditional mean as possible by applying the added control approach and checking whether the estimations are robust to a range of sample alteration in order to see if the coefficients of interest change due to omitted variable bias. These sensitivity analyses are presented later in Section 6.

Additionally, we consider the possibility of reverse causality. This is less of a concern for inputs as it is widely accepted in the literature that agricultural inputs may be regarded as predetermined vis-à-vis the level of output (Griliches, 1963; Dinar et al., 2007). This is due to the fact that agricultural production takes time to be completed and inputs are applied at the beginning of the season while the corresponding output is harvested at the end. It seems reasonable to assume that there is no correlation between the stochastic error and the predetermined inputs.

4 Data

This study uses data from the Third Integrated Household Survey (IHS3), collected from March 2010 to March 2011 by the Malawi National Statistical Office, with support from the World Bank Living Standards Measurement Study - Integrated Surveys on Agriculture (LSMS-ISA) project. The IHS3 data were collected within a two-stage cluster sampling design, and are representative at the national, urban/rural, regional, and district levels, covering 12,271 households in 768 enumeration areas (EAs). The IHS3 instruments included Household, Agriculture, Fishery, and Community Questionnaires.

All sample households were administered the multi-topic Household Questionnaire that collected individual-disaggregated information on demographics, education, health, wage employment, nonfarm enterprises, anthropometrics, and control of income from off-farm income sources, as well as data on housing, food consumption, food and non-food expenditures, food security, and durable and agricultural asset ownership, among other topics.

The sample households that were involved in agricultural activities (through ownership and/or cultivation of land, and/or ownership of livestock) were administered the Agriculture Questionnaire. The Agriculture Questionnaire solicited information on land areas, physical characteristics, labor and non-labor input use, and crop cultivation and production at the plot level, separately for the reference rainy and dry seasons.[2] The data allow for agricultural production estimates at the plot level and for the identification of the manager of the plot[3], as well as household members that owned[4] and/or worked on

---

[2] A plot was defined as a continuous piece of land on which a unique crop or a mixture of crops is grown, under a uniform, consistent crop management system, not split by a path of more than one meter in width. Plot boundaries were defined in accordance with the crops grown and the operator.

[3] For each plot, the following question was asked to identify the primary decision maker/manager: "Who in this household makes the decisions concerning crops to be planted, input use and the timing of cropping activities on this plot?" The questionnaire allowed for identification of one manager per plot, on whom individual-level information could be recovered from the Household Questionnaire.

[4] For each plot, the following question was asked to identify the plot owners: "Who owns this plot?" The question allowed up to 2 household members to be specified as owners.

each plot. Handheld global positioning system (GPS)-based locations and land areas of the plots were recorded, permitting us to link household- and plot-level data to outside geographic information system (GIS) databases.

5 Results

Using the nationally representative dataset of agricultural households in Malawi described in the preceding section, we complete the following tasks in this section: (i) we provide empirical evidence for the propositions from the theoretical model, and (ii) we decompose the gender gap in agricultural productivity into the four effects – the labor market effect, the purchased inputs effect, the endowment effect and the pure marginal productivity effect.

5.1 Description of Gender Differences

The descriptive statistics and the results from the tests of mean differences sorted by the head of household's gender are presented in Table 1. The full sample consists of 14,204 plots managed by the head of household, 23 percent of them are female headed households. The sample has been restricted to include only plots that are managed by the head of household and in which the manager works at least one hour per day on the plot. It is important to note that female headed households include mostly divorced, separated, single, or never married women, which together account for 86 percent of all female headed households. The remaining 13 percent of female headed households are married.

Thus our examination of gender differences mostly pertains to female headed households in which a male spouse is absent.

The agricultural productivity variable is proxied by the plot level gross value of output in Malawi Kwacha (MK) calculated by first multiplying the kilogram-equivalent quantity of production for each crop on a given plot by the median crop sales value per kilogram within the corresponding enumeration area (EA), and then aggregating across values of crop production. The median crop sales value per kilogram is computed within the corresponding EA only if at least 10 values are available from the survey data. Otherwise, the median crop sales value per kilogram is computed at a higher level, in the order of traditional authority, district, region, and country. Our outcome variable is computed by normalizing plot-level gross value of output with the area of the plot (land productivity). The measure of labor productivity is calculated by normalizing the plot-level gross value of output with the total number of hours of managerial labor on the plot.

Table 1 provides evidence of the gender gap: the average of the log gross value of output per hour of managerial labor is 44 percent lower for female plots, while the average of the log of gross value of output per hectare is 25 percent lower for female plots. This result provides some support for our model which predicts a gender gap in labor productivity that is larger than the gender gap in land productivity. The reasons posited for this difference in the theoretical model include labor market imperfection spillovers that lead to female-headed households allocating more labor to the farm thus decreasing labor productivity but increasing land productivity. The gender differences in agricultural productivity are also evident in the comparison of the Kernel density estimates of the log of land productivity for male and female plots, as displayed in Figure 2. The Kernel density estimates for the log of labor productivity for male and female plots are displayed in Figure 3.

The overwhelming majority of the differences in the average values of the observable covariates across male- vs. female plots in Table 1 are statistically significant at the 1 percent level. Female plots are, on average, overseen by individuals that are 8 years older and have 2.5 less years of schooling with respect to their male counterpart. A significantly higher percentage of female plots exhibit manager-owner correspondence than male plots (78 vs. 58 percent).[5] The incidences of joint ownership and exclusive-male ownership stand both at 2 percent and 3 percent respectively, among female plots. In comparison, male plots are distributed more evenly across the ownership categories of exclusive-male (39 percent), exclusive-female (22 percent), and joint male-female (19 percent).

The average GPS-based plot area for female plots is 0.34 hectare, 8 percent smaller than male plots. The use of inorganic fertilizer per hectare is in average 18 percent lower on female plots than on male plots.

In terms of household labor use, the dynamics are also highly different on female plots *vis-à-vis* their male comparators, as can be seen in Table 1. The average amount of

[5] The overwhelming majority of the owned plots (81 percent) are acquired through inheritance. Another 12 percent is reported to have been granted by local leaders. The remaining are acquired as bride price (2 percent), purchased with title (1 percent) and purchased without title (2 percent).

hours of managerial labor per hectare is 22 percent higher on female plots than on male plots, while female plots have 34 percent less family labor than male plots.[6] Female plots have higher levels of exchange labor use, while hired labor is not significantly different between male and female plots. Female plots are also 5 percentage points less likely to be associated with households that receive agricultural extension service on topics that relate to crop production and marketing. Female headed households are 20 percent less likely to participate in off-farm work than male-headed households. Lastly, male-headed households are, on average, more likely to be associated with higher levels of wealth and access to agricultural implements.[7]

5.2 Base Regression Results

Table 2 presents plot-level land productivity regression results for the male and female plot samples in columns 1 and 2 respectively. We include the explanatory variables that have been widely used in the literature (López 1984, López 1986, Peterman

---

[6] The plot-level measures of household labor input are the summations of rainy season labor hours across household members reported to have worked on a given plot. Individual labor input is computed as the multiplication of the number of weeks a household member worked on a given plot during the reference rainy season, the typical number of days worked per week during the reported number of weeks, and the typical number of hours worked per day during the reported number of weeks. The plot-level measure of hired labor (exchange) input is the sum of aggregate men, women, and child hired (exchange) labor days.

[7] The household wealth index is constructed using principal component analysis, and takes into account the number of rooms in the dwelling, a set of dummy variables accounting for the ownership of (i) dwelling, (ii) mortar, (ii) bed, (iii) table, (iv) chair, (v) fan, (vi) radio, (vii) tape/CD player, (viii) TV/VCR, (ix) sewing machine, (x) paraffin/ kerosene/ electric/ gas stove, (xi) refrigerator, (xii) bicycle, (xiii) car/motorcycle/minibus/lorry, (xiv) beer brewing drum, (xv) sofa, (xvi) coffee table, (xvii) cupboard, (xviii) lantern, (xix) clock, (xx) iron, (xxi) computer, (xxii) fixed phone line, (xxiii) cell phone, (xxiv) satellite dish, (xxv) air-conditioner, (xxvi) washing machine, (xxvii) generator, (xxviii) solar panel, (xxix) desk, and a vector of dummy variables capturing access to improved (i) outer walls, (ii) roof, (iii) floor, (iv) toilet, and (v) water source. The household agricultural implement access index is also computed using principal components analysis, and covers a range of dummy variables on the ownership of (i) hand hoe, (ii) slasher, (iii) axe, (iv) sprayer, (v) panga knife, (vi) sickle, (vii) treadle pump, (viii) watering can, (ix) ox cart, (x) ox plough, (xi) tractor, (xii) tractor plough, (xiii) ridger, (xiv) cultivator, (xv) generator, (xvi) motorized pump, (xvii) grain mail, (xviii) chicken house, (xix) livestock kraal, (xx) poultry kraal, (xxi) storage house, (xxii) granary, (xxiii) barn, and (xxiv) pig sty.

et al. 2011) including plot area, labor and non-labor inputs, plot characteristics, and farm attributes including type of crop and the presence of inter-cropping which are expected to have a direct effect on productivity. We also include covariates capturing manager and household characteristics which may affect agricultural productivity if the consumption and production decisions of the household are non-separable. Additionally, we include district level effects to account for time-invariant omitted variables at the district level.

The land productivity regressions' results are as expected; labor and non-labor inputs contribute positively to agricultural land productivity; while the area of the plot has a negative sign that reflects decreasing returns to scale of the production function. Additionally, household and manager characteristics are significant, possibly reflecting the non-separability nature of the optimization problem caused by the liquidity constraints and the labor market imperfections faced by households, as explained in the theoretical model in section 2.

We now turn to the effect of different covariates on the production of male and female plots. Table 2 shows that only six coefficients are significantly different between the male and female regressions, at the 10 percent level of significance. The coefficients that are significantly different are inorganic fertilizer, area of the plot and area squared, exchange labor, child dependency ratio and extension.

Plot area has a negative coefficient that is statistically significant at the 1 percent level in the male and female samples; however the coefficient in the female regression is more than twice as high as in the male regression. We calculate the marginal productivity of land, which is shown in table 3 and find that it is 12 percent lower in female managed plots compared to male-managed plots.

The log of inorganic fertilizer use per hectare is positively associated with the log of gross value of output per hectare, irrespective of the plot sample. However, the return to inorganic fertilizer use (i.e. the coefficient of inorganic fertilizer) is higher within the male plot sample than within the female.

The log of managerial labor has a positive coefficient that is statistically significant at the 1 percent level across the male and female samples. As shown in table 3, the marginal productivity of female managers is 37 percent lower than the labor productivity of male managers. The coefficient of household labor is positive and significant in the female and male plots, but larger on male plots, while the coefficient of exchange labor is only significant on the male plot sample.

The child dependency ratio, which is defined as the number of household members below the age of 10 divided by the number of household members aged 10 years and above, has a substantial negative coefficient that is statistically significant at the 1 percent level only within the female-managed plot sample. The comparable statistics for the male-managed plot sample is negative and statistically insignificant.

With respect to the household characteristics, household size has a positive coefficient that is statistically significant irrespective of the plot sample; the magnitude of the coefficient within the female plot sample is larger than within the male-managed plot sample. The gender differences in returns to child dependency ratio after controlling for household size imply that the burden of childcare is more likely to reduce female agricultural productivity than male agricultural productivity. The distance of the plot to the household compound is negative and statistically insignificant for the male and female samples, while the distance of the household to the nearest road is negative and not statistically significant for the female sample alone.

5.3 Linking To The Theoretical Model

Our theoretical model provides three key features: (i) the existence of non-separability of consumption and production decisions by the household due to the presence of market failures, (ii) female-managed plots use fewer inputs due to labor market imperfections and the gender specific off-farm time burden incurred when engaging in off-farm work that result in less off-farm labor income and (iii) given that female headed households have lower access to non-labor inputs, female plots have lower agricultural labor productivity in comparison to male plots.

As the base regressions in table 2 indicate, household and manager characteristics are significant, possibly reflecting the effect of the labor and liquidity constraints faced by the household over the production decisions. This provides support for (i) a key assumption in the model that the nature of the optimization problem is non-separable. The descriptive statistics show the presence of a gender gap, lending support to point (iii). The average gross value of output per area of the plot is 25 percent lower for female plots than male plots, and the average gross value of output per hour of managerial labor is 44 percent lower for female plots than male plots.

Base regression results in table 2 further illustrate that the child dependency ratio has a significant negative effect for female-managed plots but is insignificant for male-managed plots even after controlling for household size. This implies that the burden of childcare is more likely to reduce female agricultural productivity than male agricultural productivity. In support of (ii) we find that female headed households are 20 percent less likely to participate in off-farm activities than male headed households and that female managers on average work 22 percent more hours per hectare than male managers, while the use of inorganic fertilizer is 20 percent lower in female plots.

Finally, in Table 3 we present the estimates of the marginal land and labor productivities of male and female managers derived from the regressions in Table 2. The marginal productivity of land is 12 percent lower for female managers compared to male managers, while the marginal labor productivity of female managers is 37 percent lower than the marginal labor productivity of male managers. Consistent with our theoretical model, we provide the following explanations for the gender differences in land and labor productivity. The gender gap in marginal land productivity is due to lower access to non-labor inputs for female managed plots relative to men given the complementarities between labor and non-labor inputs. The gender gap in marginal labor productivity is due to increased allocation of labor in female managed plots to farm activities given the labor market discrimination that provides disincentives for female head of households to engage in off-farm work.

5.4 Mean Decomposition

In the decomposition methodology, we uncovered four components of gender differences in agricultural productivity – labor market, purchased inputs, household endowment and pure marginal productivity effects outlined in section 3. We decompose the mean gender gap in agricultural land productivity and the agricultural labor productivity in order to rank the importance of each of the four components as indicated in (15) and (A.4) respectively. The decomposition uses the base regressions (section 5.2) which correspond to equations (12a) and (12b). We find that the decomposition results are consistent with the theoretical model. Gender gaps exist in favour of male-headed households for both labor and land productivity, however the gender differences are far greater for labor productivity.

5.4.1 Land Productivity Decomposition

As mentioned in the theoretical model, we predict that the agricultural productivity of female headed households will be lower when compared to male managed plots, the theoretical model also states that the gender gap in agricultural productivity is explained by how differently households are affected by liquidity and labor market constraints which are influenced by the composition of the household and its preferences. From the theoretical model we expect the gender gap to be decomposed into differences in levels of use of purchased inputs (male managed plots will have higher use), managerial labor (higher levels for female-managed plots) and differences in levels of the assets and characteristics of the household.

Table 4 shows the decomposition of the mean gender differential in agricultural land productivity into the four effects specified in equation (15). The gender gap in land productivity is estimated at 25 percent. The four effects are all significant at the 1 percent level.

The decomposition indicates that the purchased inputs and household endowment effects account for 57 and 37 percent of the total land gender gap respectively. According to our model, these two effects are a manifestation of the impact that liquidity constraints have on households as well as labor market imperfections, which are different for male and female headed households. Differences in purchased inputs reflect the direct effect of liquidity constraints and the indirect effect of labor market imperfections, while differences in the endowment of the household may be a result of the long term impact of such constraints. The labor market effect is estimated to be -0.5 percentage points, which represents 19 percent of the gender gap. The latter means that if women worked the same number of hours as men, their land productivity would be lower and the gender gap in land productivity would rise to 30 percent. The labor, endowment and purchased inputs effects account for 70 percent of the gender gap in land productivity. The remaining 30 percent of the gender gap in agricultural land productivity is explained by differences in coefficients (the pure marginal productivity effect), which includes the differential effect of the child dependency ratio, distance of the plot to the household and distance to the closest market, all of which are proxies to the gender specific off-farm time burden due to engagement in off-farm activities.

5.4.2 Labor Productivity Decomposition

The theoretical model predicts specifically that labor productivity will be lower in female-headed households due to the different impact that the liquidity constraints, labor market discrimination and gender specific off-farm time burden have over them compared to male-headed households (Equation (11)). As can be seen in Table 5, these predictions are consistent with the empirical results.

The gender gap in agricultural labor productivity is estimated at 44 percent. The agricultural labor productivity gender gap is larger than the agricultural land productivity gender gap, due to the labor market effect which accounts for 34 percent of the total gap. In this case the fact that female managers work more hours in the plot than male managers exacerbates the agricultural labor productivity gender gap, unlike its effect on land productivity, where it reduces the gap. The labor market effect may be attributed in part to the direct effect of gender differential treatment in the labor market and should be considered as an upper bound.

The purchased inputs and the household endowment effects are 29 and 20 percent of the gender gap respectively, which together with the labor effect explain 83 percent of the gender gap in agricultural labor productivity; the remaining 17 percent is explained by the pure marginal productivity effect.

5.5 Unconditional Decile Decomposition

So far we have performed the decomposition at the mean. However going beyond the “average” farmer and understanding the *heterogeneity* in constraints faced by farmers with different gender and productivity profiles is crucial for the design and implementation of better targeted interventions aimed at bridging the gender gap. An important question is whether our findings, which are based on the sample means, are robust to the decomposition of alternative distributional statistics beyond the mean.

A method that is similar in spirit to the mean decomposition uses the recentered influence function (RIF) regressions proposed by Firpo et al. (2009) and provides a straightforward framework within which across-group differences in any distributional statistic could be decomposed. We rely on the RIF decomposition to provide estimates of the decomposition of the gender gap at different deciles of the agricultural productivity distribution. A detailed description of the methodology is presented in Appendix B.

5.5.1 Land Productivity Unconditional Decile Decomposition

Table 6 presents the gender gap estimates and RIF decompositions, both at the mean and at each decile of the agricultural land productivity distribution. The graphical representation of these findings is reported in Figure 5. The four effects in figure 5 are based on the RIF regressions that use the same set of independent variables included in the base specification for the mean decomposition.

Two key findings emerge from Table 6. First, the magnitude of the gender gap and the share of the gender gap attributed to the effect of purchased inputs increase steadily along the agricultural land productivity distribution. The mean and median gender gap is estimated to be 25 and 21 percent respectively. The magnitudes of the gender gap range between 17 percent at the 10^{th} percentile to 39 percent at the 90^{th} percentile. The purchased inputs effect of the decomposition accounts for 52 percent of the gender gap at the mean and 61 percent of the gender gap at the 90th percentile. This implies that the purchased inputs effect is the largest contributor to the gender gap.

Second, the household endowment effect declines along the land productivity distribution. The decline of the household endowment effect on the gender gap is from 65 percent at the 10^{th} percentile to 35 percent at the 90^{th} percentile.

One interpretation of the declining importance of the household endowment effect and the increasing importance of the purchased inputs effect towards the land productivity gender gap is that at the lower end of the productivity distribution female-headed households tend to be relatively more deprived of endowments than male-headed households. However, at the other end of the productivity distribution, the relative deprivation of endowments is less important for females than males, given that the households overall tend to be richer. Thus, the purchased inputs effect becomes the dominant factor in explaining the gender productivity gap as differential access to markets by gender become more important.

5.5.2 Labor Productivity Unconditional Decile Decomposition

Table 7 presents the gender gap estimates and RIF decompositions at each decile and the mean of the agricultural labor productivity distribution. The graphical representation of these findings is reported in Figure 6.

The estimates of the gender gap increase steadily across the labor productivity distribution. The mean and median gender gap is estimated to be 44 and 49 percent respectively. The magnitudes of the gender gap range between 35 percent at the 10th percentile to 52 percent at the 80th percentile. However, at the 90th percentile the gender gap declines to 40 percent, which is still a high and significant value. This is a surprising result and could be due to increasing efficiency of women, despite having less access to purchased inputs and lower endowment than men.[8]

The labor market and the purchased input effects on the gender gap increase steadily along the productivity distribution. The increase of the labor effect on the gender gap is from 34 percent at the 10th percentile to 40 percent at the 80th percentile. The increase of the purchased input effect on the gender gap is from 23 percent at the 10th percentile to 38 percent at the 80th percentile. The two effects – labor market and

[8]We estimated our regressions dropping the top 10 percent of the productivity and the mean results do not change dramatically. The normalized difference between the male and female samples of the top 10 percent of the land productivity distribution was calculated. Only 8 out of 26 independent variables and 30 district dummies have a normalized difference greater than 0.25, thus, it is possible to conclude that there is overlapping support across the groups at the higher end of the distribution.

purchased input – tend to work in the same direction due to the complementarity between labor and non-labor inputs.

6 Sensitivity Analyses

We are concerned about two main issues: (i) the validity of the decomposition methodology employed and (ii) whether the gender differences of interest – labor, purchased inputs, endowment including child dependency ratio, plot and household locations, - are robust to various specifications. As noted earlier, the crucial assumptions for the validity of the aggregate decomposition include overlapping support and ignorability. The key assumptions additionally required by the detailed decomposition are additive linearity and zero conditional mean. A methodology that is proposed by Imbens and Rubin (2009) to assess the feasibility of the overlapping support assumption is centered on the idea of calculating a scale-free normalized difference for each covariate. They assert that the overlapping support across the groups of interest, in our case female vs. male plots, is adequate if the scale-free normalized differences across most of the covariates are less than 0.25. Table C.1 in the Appendix presents the scale-free normalized difference of the variables used in the regressions. Only 5 out of 26 independent variables (and 30 district dummies) have a normalized difference greater than 0.25.

In trying to lend support to ignorability and zero conditional mean assumptions, we use all available data and econometric tools at our disposal, and first rely on an empirical approach that was pioneered by Altonji (1988), Murphy and Topel (1990), and

Altonji et al. (2005), based on the idea that the amount of selection on observable variables provides a guide to the extent of selection on unobservable counterparts. We use an informal version of the methodology applied by Acemoglu et al. (2001), and incorporate into our base specification, in a phased-in fashion, thematically-grouped control variables such that each regression is estimated with a different set of additional independent variables and that the results are compared to those from the base specification. Our purpose is to gauge the stability of the key regression coefficients that underlie our decomposition results. If the coefficients on the covariates included in the base specification are stable subsequent to the incorporation of additional covariates, they are less likely to change if we are able to take into account potentially missing omitted variables.

To perform this analysis, we consider the following sets of variables: (i) enumeration area effects, (ii) plot geospatial characteristics informed by GIS data, (iii) other plot characteristics solicited by the IHS3, (iv) additional household characteristics, and (v) additional community characteristics. Table C.2 in the Appendix includes the detailed list of the variables included in each set. Tables C.3 and C.4 present the base regression results and the estimates from the regressions including the additional controls for the male and female plot samples, respectively. An overwhelming majority of the coefficients, with respect to the base specification, are stable across the specifications and the plot samples, and do not change sign or significance. This suggests that the assumptions of ignorability and zero conditional mean might not be unfounded.

In addition, we checked for extreme data points that may dominate the sign and significance of key estimates. We conducted two types of dominance tests. In order to account for extreme data points, we first re-estimated the model by excluding observations in the top and bottom 1 percent of the land productivity. The same procedure is followed by re-estimating the model without observations in the top and bottom 1 percent of managerial labor. The parameters are robust to the sample changes. Signs, significance and magnitudes of the parameter estimates from these models are shown in Table C.5 in the Appendix.

7 Conclusion

This study presents a theoretical model that sheds light on the mechanisms underlying gender differences in agricultural productivity. It focuses on the effects of labor and credit market imperfections as well as on other long term structural factors that affect female- and male-headed households differently. The empirical approach provides evidence of the relative quantitative importance of factors that lie behind the gender gap in agricultural labor productivity, providing evidence consistent with the predictions of the theoretical model.

This study theoretically and empirically uncovers the importance of market imperfections behind the gender gap in agricultural labor productivity. Liquidity constraints, labor market discrimination and off-farm time burden due to engagement in off-farm work, which differ greatly between men and women, result in lower agricultural labor productivity in plots belonging to female- in comparison to male-headed households.

Our paper proposes several possible mechanisms consistent with a gender gap in agricultural labor productivity. Households diminish the effect of credit constraints on their production decisions via participation in off-farm activities. Participation in off-farm activities depends on access to off-farm opportunities and the composition of the household. Female-headed households may be discouraged to participate in off-farm activities due to discrimination. Additionally, female-headed households might opt not to participate in off-farm work after considering the additional gender specific off-farm time burden, which is the opportunity cost of carrying out traditional household activities attributed to them (such as child care, getting water, and cooking). The lack of participation in off-farm work by women may imply they have lower access to productive inputs, exacerbating the effect of liquidity constraints and resulting in them being less productive than men. It is not only the effect of the labor and credit market discrimination but also the synergy between them that disfavors women ultimately leading to their lower agricultural labor productivity in comparison to men. The empirical estimation is consistent with the predictions of the model, the effect of labor market imperfections, liquidity constraints and household characteristics lead to lower agricultural labor productivity on plots belonging to female-headed households.

REFERENCES

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). "The colonial origins of comparative development: An empirical investigation." *American Economic Review*, 91 (5), 1369-1401.

Alene, A. D., Manyong, V. M., Omanya, G. O., Mignouna, H. D., Bokanga, M., & Odhiambo, G. D. (2008). "Economic efficiency and supply response of women as farm managers: Comparative evidence from Western Kenya." *World Development* 36 (7), 1247–1260.

Akresh, R. (2005). "Understanding pareto inefficient intrahousehold allocations." Discussion Paper No. 1858, IZA. Retrieved from http://ftp.iza.org/dp1858.pdf.

Altonji, J. (1988). "The effects of family background and school characteristics on education and labor market outcomes." Manuscript, Department of Economics, Northwestern University.

Altonji, J., Elder, T., & Taber, C. (2005). "Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools." *Journal of Political Economy*, 113, 151-184.

Barrett, C.B. (1996), "On price risk and the inverse farm size-productivity relationship", Journal of Development Economics 51:193–215.

Barrett, C.B., S.M. Sherlund, and A.A. Adesina (2008), Shadow wages, allocative inefficiency, and labor supply in smallholder agriculture. *Agricultural Economics* 38(2008): 21-34.

Binswanger, H., M. Rosenzweig, eds. (1986): Contractual Arrangements, Employment and Wages in Rural Labor Markets in Asia. New Haven: Yale University Press

Blinder, A. (1973). "Wage discrimination: reduced form and structural estimates." *Journal of Human Resources*, 8, 436–455.

Dinar, A. , Karagiannis, G. and Tzouvelekas, V. (2007) "Evaluating the impact of agricultural extension on farm's performance in Crete: a nonneutral stochastic frontier approach." Agricultural Economics, vol. 36, issue 2, 135-146.

Doepke, M., & Tertilt, M. (2011). "Does female empowerment promote economic development?" Discussion Paper No. DP8441, CEPR.

Doss, C. (2011). "If women hold up half the sky, how much of the world's food do they produce?" FAO- ESA Working Paper No. 11-02. Retrieved from http://www.fao.org/docrep/013/am309e/am309e00.pdf

Doss, C., Raney, T., Anriquez, G., Croppenstedt, A., Gerosa, S., Lowder, S., Matuscke, I. & Skoet, J. (2011). "The role of women in agriculture." FAO-ESA Working Paper No. 11-02. Retrieved from http://www.fao.org/docrep/013/am307e/am307e00.pdf

Eswaran, M., and Kotwal, A. (1986)."Access to Capital and Agrarian Production Organization."
Econ. J. 96(June 1986):482–98.

FAO (1984). *Women in food production and food Security in Africa*. Rome, Italy: FAO.

FAO (2011). T*he state of food and agriculture: Women in agriculture – closing the gender gap for development.* Rome, Italy: FAO. Retrieved from http://www.fao.org/docrep/013/i2050e/i2050e.pdf.

FAO (2011b). *Gender Inequalities in Rural Employment in Malawi, An Overview*. Rome, Italy: FAO. Retrieved from http://www.fao.org/docrep/016/ap092e/ap092e00.pdf.

Firpo, S., Fortin, N. M., & Lemieux, T. (2009). "Unconditional quantile regressions." *Econometrica*, 77 (3), 953–973.

Fortin, N. M. (2006). "Greed, altruism, and the gender wage gap." Working Paper, Department of Economics, University of British Columbia. Retrieved from http://faculty.arts.ubc.ca/nfortin/Fortinat8.pdf.

Fortin, N., T. Lemieux, S. Firpo (2011). "Decomposition methods." In O. Ashenfelter & D. Card (Eds.), *Handbook of labor economics* (Vol. 4, Part A, 1-102) Amsterdam, Netherlands: North-Holland.

Gilbert, R. A., Sakala, W. D., & Benson, T. D. (2002). "Gender analysis of a nationwide cropping system trial survey in Malawi." *African Studies Quarterly*, 6 (1), Retrieved from http://web.africa.ufl.edu/asq/v6/v6i1a9.htm.

Goldstein, M., & Udry, C. (2008). "The profits of power: Land rights and agricultural investment in Ghana." *Journal of Political Economy* 116 (6), 981–1022.

Griliches, Z. (1963). "Estimation of the aggregate production function from cross sectional data. J Farm Econ. 45, 419-428

Hertz, T., P. Winters, A.P. De La O, E.J. Quinones, C. Azzari, B. Davis and A. Zezza (2009). "Wage inequality in international perspective: effects of location, sector, and gender." Paper presented at the FAO-IFAD-ILO Workshop on Gaps, trends and current research in gender dimensions of agricultural and rural employment: differentiated pathways out of poverty Rome, 31 March - 2 April 2009, FAO, Italy.

Imbens, W.G., & Rubin, B.D. (2009). *Causal inference in statistics, and in the social and biomedical sciences*. New York, NY: Cambridge University Press.

Irz, X., Lin, L., Thirtle, C., & Wiggins, S. (2001). "Agricultural productivity growth and poverty alleviation." *Development Policy Review*, 19 (4), 449-466.

López, R. E., (1984). "Estimating labor supply and production decisions of self-employed farm producers," European Economic Review, Elsevier, vol. 24(1), pages 61-82.

López, R. (1986). "Structural models of the farm household that allow for interdependent utility and profit maximization decisions." In Inderjit J. Singh, Lyn Squire, and John Strauss (eds), *Agricultural Household Models-Extensions, Applications and Policy*. Baltimore: The Johns Hopkins University Press..

Moock, P. R. (1976). "The efficiency of women as farm managers: Kenya." *American Journal of Agricultural Economics*, 58 (5), 831–835.

Murphy, K. M., & Topel, R. H. (1990). "Efficiency wages reconsidered: Theory and evidence." In Y. Weiss & R. H. Topel (Eds.), *Advances in the theory and measurement of unemployment* (204-240). New York, NY: St. Martin's Press.

Oaxaca, R. (1973). "Male-female wage differentials in urban labor markets." *International Economic Review*, 14, 693–709.

Oladeebo, J. O., & Fajuyigbe, A. A. (2007). "Technical efficiency of men and women upland rice farmers in Osun State, Nigeria." *Journal of Human Ecology*, 22 (2), 93–100.

O'Neill, J. E., & O'Neill, D. M. (2006). "What do wage differentials tell about labor market discrimination?" In S. W. Polachek, C. Chiswick & H. Rapoport (Eds.) *The economics of immigration and social diversity* (Research in labor economics series, Vol. 24, 293-357). Amsterdam, Netherlands: Elsevier.

Peterman, A., Behrman, J. & Quisumbing, A. (2010). " A review of Empirical Evidence on Gender Differences in nonland agricultural inputs, technology and services in developing countries. IFPRI Discussion Paper 00975

Peterman, A., Quisumbing, A., Behrman, J., & Nkonya, E. (2011). "Understanding the complexities surrounding gender differences in agricultural productivity in Nigeria and Uganda." *Journal of Development Studies*, 47 (10), 1482-1509.

Quisumbing, A., Payongayong, E., Aidoo, J. B., & Otsuka, K. (2001). “Women’s land rights in the transition to individualized ownership: Implications for the management of tree resources in western Ghana.” *Economic Development and Cultural Change*, 50 (1), 157–182.

Saito, K. A., Mekonnen, H., & Spurling, D. (1994). “Raising the productivity of women farmers in sub-Saharan Africa.” Discussion Paper Series 230, Africa Technical Department, The World Bank. Retrieved from http://elibrary.worldbank.org/content/book/9780821327494.

Tiruneh, A., Testfaye, T., Mwangi, W., & Verkuijl, H. (2001). “Gender differentials in agricultural production and decision-making among smallholders in Ada, Lume, and Gimbichu woredas of the central highlands of Ethiopia.” International Maize and Wheat Improvement Center and Ethiopian Agricultural Research Organization: Mexico, DF, and Addis Ababa. Retrieved from http://repository.cimmyt.org/xmlui/bitstream/handle/10883/1018/73252.pdf?sequence=1

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

United Nations. Economic Commission for Africa. Human Resources Development Division.UNECA (1982) “Women: the neglected human resource for African Development.” Canadian Journal of African Studies, 6 (2), 359-370.

Vargas Hill, R., & Vigneri, M. (2011). “Mainstreaming gender sensitivity in cash crop markets supply chains.” Working Paper No. 11-08, ESA, FAO. Retrieved from http://www.fao.org/docrep/013/am313e/am313e00.pdf.

World Bank (WB) (2011). W*orld development report 2012: Gender equality and development.* Washington, DC: The World Bank. Retrieved from http://siteresources.worldbank.org/INTWDR2012/Resources/7778105-1299699968583/7786210-1315936222006/Complete-Report.pdf.

Table 1: Descriptive Statistics & Results from Tests & Mean Differences by Gender of the Plot Manager

	Pooled Sample	*Male Sample*	*Female Sample*	*Difference*	
Sample size	14,204	10,962	3,242		
Outcome Variable					
Ln Value Output per hectare	10.42	10.48	10.23	0.2	***
Ln Value Output per hour of managerial labor	4.56	4.66	4.22	0.4	***
Plot Managerial Labor Input Use					
Ln Managerial Labor (hours/ha)	5.9	5.8	6.0	-0.2	***
Purchased Inputs					
Pesticide/herbicide use yes/no	0.019	0.021	0.012	0.009	***
Organic Fertilizer use yes/no	0.112	0.114	0.105	0.009	
Ln Inorganic Fertilizer (kg/ha)	3.275	3.317	3.135	0.2	***
Ln Hired labor (days/ha)	0.547	0.555	0.520	0.0	
Agricultural implements Asset Index	0.687	0.842	0.161	0.7	***
Proportion of area of the plot under improved varieties	0.372	0.383	0.333	0.0	***
Proportion of area of the plot under export crops	0.079	0.094	0.028	0.1	***
Endowment of the Household					
Ln Area of the plot (ha)	-1.224	-1.203	-1.294	0.1	***
Ln Area of the plot (ha) Squared	1.978	1.926	2.153	-0.2	***
elevation (m)	893.9	908.8	843.5	65.3	***
plot distance to hh	1.970	2.058	1.675	0.4	**
Inter-croppped	0.353	0.325	0.448	-0.1	***
Manager is equal to one of the owners	0.625	0.579	0.782	-0.2	***
Age of the manager	42.98	41.18	49.07	-7.9	***
Years of Schooling of the manager	5.206	5.796	3.214	2.6	***
Non-Managerial Household Labor (hours/ha)	531.5	577.1	377.6	199.5	***
Ln Non-Managerial Household Labor (hours/ha)	5.317	5.767	3.795	2.0	***
Ln Exchange labor (days/ha)	0.213	0.182	0.316	-0.1	***
Household Size	4.892	5.157	3.995	1.2	***
Dependency Ratio	0.704	0.701	0.713	0.0	
Ag extension services receipt	0.311	0.322	0.271	0.1	***
HH has any off-farm income	0.423	0.443	0.354	0.1	***
HH receives other transfers/safety net help	0.216	0.213	0.227	0.0	*
Wealth Index	-0.701	-0.605	-1.025	0.4	***
HH Distance (KMs) to Nearest ADMARC	8.196	8.195	8.200	0.0	

note: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table2: Base OLS Regression Results Underlying the Mean Decomposition
Dependent Variable: Ln[Plot Gross Value of Output per hectare]

	Male Managed Plot Sample	Female Managed Plot Sample	Difference in Coefficients
Labor			
Ln Managerial Labor (hours/ha)	0.227***	0.234***	
	(0.011)	(0.018)	
Purchased Inputs			
Pesticide/herbicide use yes/no	0.446***	0.618***	
	(0.054)	(0.121)	
Organic fertilizer use yes/no	0.021	0.036	
	(0.024)	(0.044)	
Ln Inorganic Fertilizer (kg/ha)	0.070***	0.058***	***
	(0.003)	(0.006)	
Ln Hired labor (days/ha)	0.102***	0.113***	
	(0.007)	(0.012)	
Agricultural Implements Access Index	0.036***	0.048***	
	(0.006)	(0.012)	
Proportion of Plot Area Under Improved Seeds	0.041**	0.025	
	(0.018)	(0.034)	
Proportion of Plot Area Under Export Crops	1.078***	1.132***	
	(0.030)	(0.086)	
Household Characteristics and Endowment			
Ln GPS Total Area of the plot (ha)	-0.155***	-0.354***	**
	(0.040)	(0.074)	
Ln GPS Total Area of the plot (ha) Squared	0.056***	0.006	*
	(0.014)	(0.025)	
Elevation (m)	0.000**	0.000**	
	(0.000)	(0.000)	
Plot distance to household	-0.000	-0.000	
	(0.001)	(0.002)	
Intercropped	0.230***	0.273***	
	(0.022)	(0.037)	
Manager is equal to one of the owners	-0.007	-0.002	
	(0.016)	(0.033)	
Age of the manager	-0.001**	-0.001	
	(0.001)	(0.001)	
Years of Schooling of the manager	0.003	0.011**	
	(0.002)	(0.005)	
Ln Non-Managerial Household Labor (hours/ha)	0.015**	0.012**	
	(0.006)	(0.005)	
Ln Exchange labor (days/ha)	0.042***	0.013	*
	(0.011)	(0.015)	

Table2: Base OLS Regression Results Underlying the Mean Decomposition (Cont'd)
Dependent Variable: Ln[Plot Gross Value of Output per hectare]

	Male Managed Plot Sample	Female Managed Plot Sample	Difference in Coefficients
Household Size	0.013***	0.024***	
	(0.004)	(0.008)	
Dependency Ratio	-0.008	-0.068***	**
	(0.016)	(0.018)	
Agricultural Extension Receipt	0.030*	0.131***	***
	(0.017)	(0.031)	
HH has any off-farm income	-0.063***	-0.036	
	(0.016)	(0.029)	
HH receives other transfers/safety net help	0.007	-0.030	
	(0.020)	(0.033)	
Wealth Index	0.062***	0.060***	
	(0.005)	(0.009)	
Distance to Nearest ADMARC (KM)	0.004**	-0.000	
	(0.002)	(0.003)	
Constant	7.958***	8.307***	
	(0.261)	(0.757)	
Number of observations	10,962	3,242	
R2	0.380	0.373	
Adjusted R2	0.376	0.362	

Robust standard errors in parenthesis.
note: *** p<0.01, ** p<0.05, * p<0.1

Table 3. Marginal Productivity of land and labor in Agricultural Land Productivity

Marginal Productivity of Land in Agricultural Production		
Male Headed Households	Female Headed Households	Difference
113,213***	81,877***	31,336***
(2,237)	(3,161)	(3,872)

Marginal Productivity of Labor in Agricultural Production		
Male Headed Households	Female Headed Households	Difference
28.4***	17.8***	10.6***
(1.38)	(1.42)	(1.98)

note: *** p<0.01, ** p<0.05, * p<0.1

Table 4. Decomposition of the Gender Gap in Agricultural Land Productivity

A. Mean Gender Gap	
Male Plots	10.477***
	(0.009)
Female Plots	10.231***
	(0.016)
Gender Gap	0.246***
	(0.019)

B. Decomposition of the Mean Gender Gap+	
Labor Market Effect	-0.046***
	(0.005)
Purchased Inputs Effect	0.129***
	(0.012)
Household Endowment Effect	0.090***
	(0.018)
Pure Marginal Productivity Effect	0.073***
	(0.023)

note: *** p<0.01, ** p<0.05, * p<0.1,
+ all the effects sum up to the Mean Gender Gap

Table 5. Decomposition of the Gender Gap in Agricultural Labor Productivity

A. Mean Gender Gap	
Male Plots	4.664***
	(0.011)
Female Plots	4.223***
	(0.019)
Mean Gender Gap	0.442***
	(0.022)
B. Decomposition of the Mean Gender Gap[+]	
Labor Market Effect	0.150***
	(0.014)
Purchased Inputs Effect	0.129***
	(0.012)
Household Endowment Effect	0.090***
	(0.018)
Pure Marginal Productivity Effect	0.073***
	(0.023)

note: *** p<0.01, ** p<0.05, * p<0.1,
+ all the effects sum up to the Mean Gender Gap

Table 6. Decomposition of the Gender Differential in Agricultural Productivity at Selected Points of the Agricultural Land Productivity Distribution

A. Mean Gender Gap	Mean	10th Percentile	20th Percentile	30th Percentile	40th Percentile	50th Percentile	60th Percentile	70th Percentile	80th Percentile	90th Percentile
Male Plots	10.48***	9.25***	9.68***	9.99***	10.24***	10.46***	10.69***	10.94***	11.26***	11.73***
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Female Plots	10.23***	9.08***	9.51***	9.78***	10.02***	10.25***	10.44***	10.68***	10.93***	11.34***
	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
Mean Gender Gap	0.25***	0.17***	0.17***	0.21***	0.22***	0.21***	0.26***	0.26***	0.33***	0.39***
	(0.02)	(0.04)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.04)
B. Decomposition of the Mean Gender Gap+										
Labor Market Effect	-0.05***	-0.07***	-0.04***	-0.03***	-0.04***	-0.04***	-0.04***	-0.04***	-0.03***	-0.05***
	(0.005)	(0.010)	(0.007)	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)	(0.005)	(0.009)
Purchased Inputs Effect	0.13***	0.09***	0.06***	0.07***	0.08***	0.10***	0.11***	0.13***	0.14***	0.24***
	(0.012)	(0.025)	(0.017)	(0.015)	(0.014)	(0.013)	(0.013)	(0.015)	(0.015)	(0.025)
Household Endowment Effect	0.09***	0.11***	0.11***	0.09***	0.09***	0.09***	0.07***	0.10***	0.09***	0.14***
	(0.018)	(0.042)	(0.030)	(0.025)	(0.024)	(0.021)	(0.021)	(0.024)	(0.023)	(0.039)
Pure Marginal Productivity Effect	0.07***	0.03***	0.05***	0.08***	0.08***	0.06***	0.11***	0.08***	0.13***	0.06***
	(0.02)	(0.05)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)
Number of observations	14,204	14,204	14,204	14,204	14,204	14,204	14,204	14,204	14,204	14,204

note: *** p<0.01, ** p<0.05, * p<0.1,
+ all the effects sum up to the Gender Gap

Table 7. Decomposition of the Gender Differential in Agricultural Productivity at Selected Points of the Agricultural Labor Productivity Distribution

A. Mean Gender Differential	Mean	10th Percentile	20th Percentile	30th Percentile	40th Percentile	50th Percentile	60th Percentile	70th Percentile	80th Percentile	90th Percentile
Male Plots	4.66***	3.25***	3.73***	4.06***	4.34***	4.63***	4.92***	5.23***	5.60***	6.15***
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Female Plots	4.22***	2.90***	3.30***	3.59***	3.87***	4.14***	4.41***	4.69***	5.08***	5.75***
	(0.02)	(0.04)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.04)
Mean Gender Gap	0.44***	0.35***	0.43***	0.47***	0.48***	0.49***	0.51***	0.54***	0.52***	0.40***
	(0.02)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.05)
B. Decomposition of the Mean Gender Differential										
Labor Market Effect	0.15***	0.12***	0.11***	0.12***	0.14***	0.15***	0.16***	0.17***	0.21***	0.24***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Purchased Inputs Effect	0.13***	0.08***	0.09***	0.09***	0.11***	0.14***	0.15***	0.18***	0.20***	0.16***
	(0.01)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
Household Endowment Effect	0.09***	0.11**	0.10***	0.12***	0.14***	0.12***	0.11***	0.07**	0.04	0.10**
	(0.02)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)
Pure Marginal Productivity Effect	0.07***	0.05	0.12***	0.14***	0.08**	0.08**	0.10***	0.13***	0.07*	-0.10
	(0.02)	(0.06)	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)	(0.04)	(0.04)	(0.06)
Number of observations	14,204	14,204	14,204	14,204	14,204	14,204	14,204	14,204	14,204	14,204

note: *** p<0.01, ** p<0.05, * p<0.1

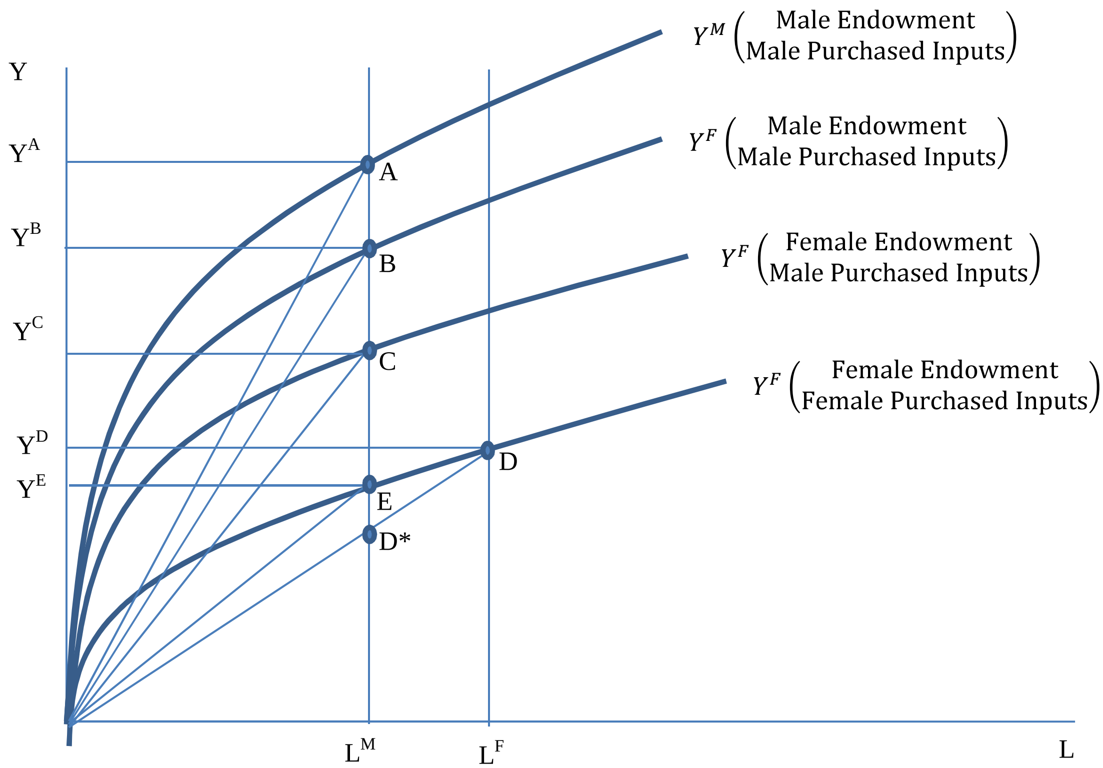


Figure 1. Land and Labor Productivity Gender Gap Decomposition into Endowment, Purchased Inputs, Labor Market and Pure Marginal Productivity Effects

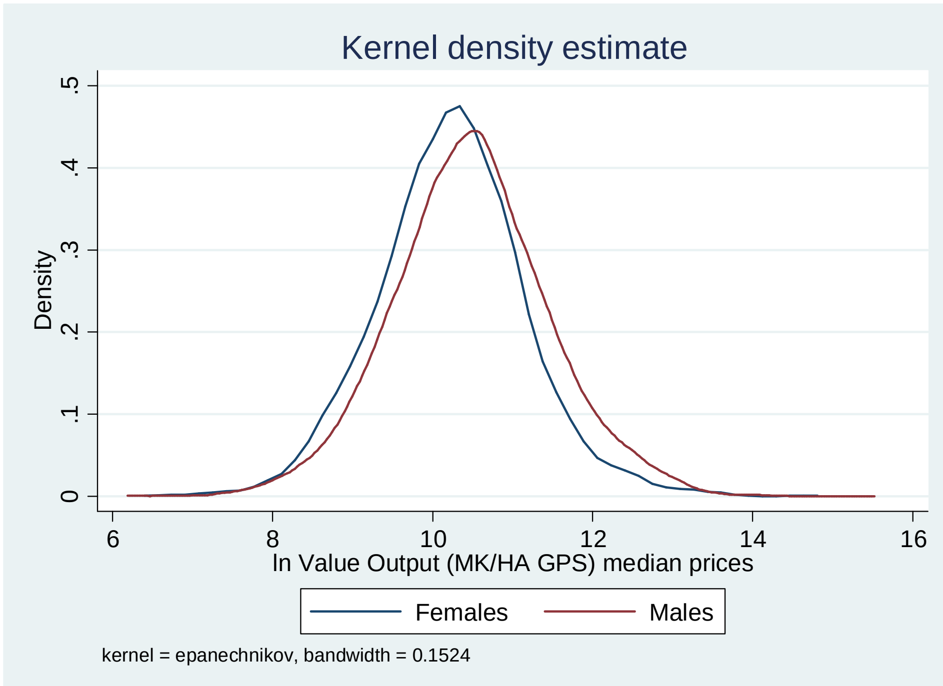


Figure 2. Kernel Density Estimates of the Log of Gross Value of Output per Hectare for Male- and Female-Managed Plot Samples

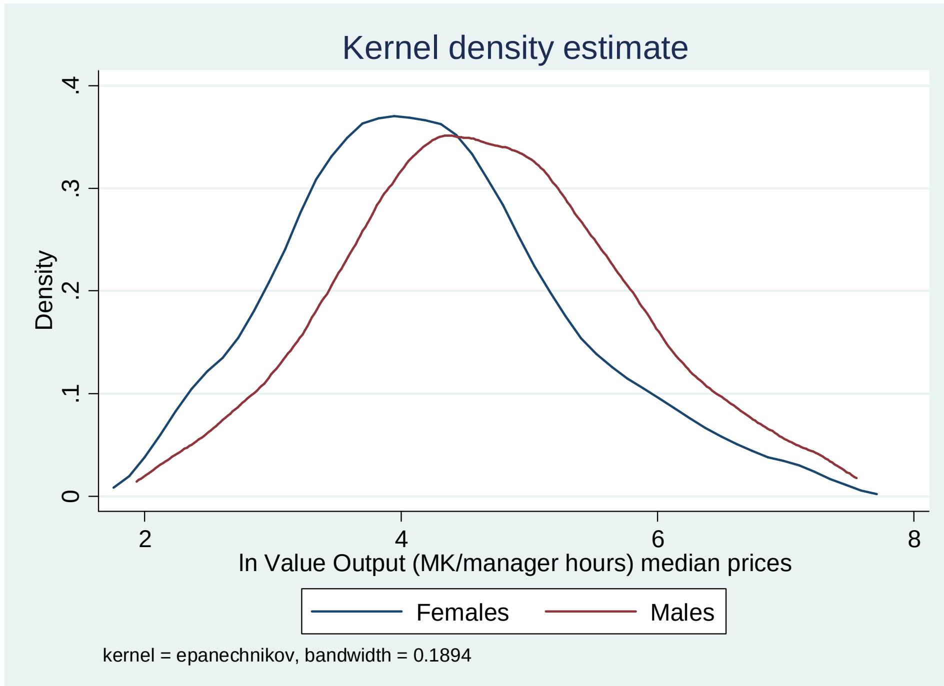


Figure 3. Kernel Density Estimates of the Log of Gross Value of Output per Managerial Labor for Male- and Female-Managed Plot Samples

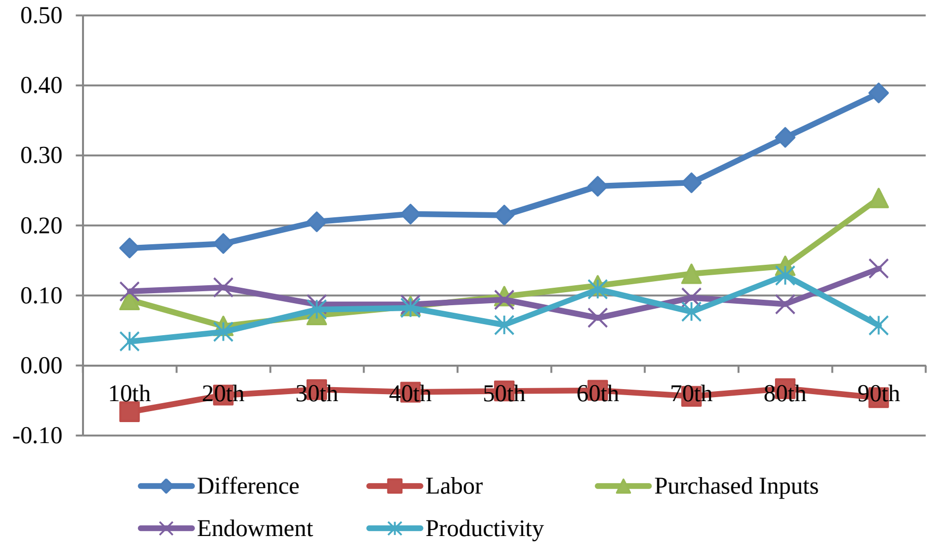


Figure 4. Decomposition of the Land Productivity Gender Gap at different deciles of the Land Productivity Distribution

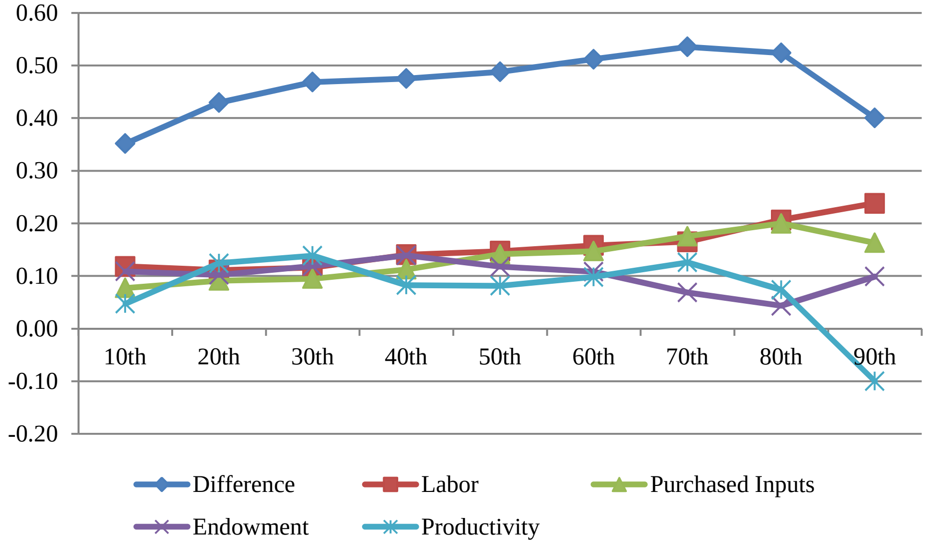


Figure 5. Decomposition of the Labor Productivity Gender Gap at different deciles of the Land Productivity Distribution

Chapter 2: Why has Europe become environmentally cleaner? Decomposing the roles of fiscal, trade and environmental policies

By Ramón López and Amparo Palacios López

Forthcoming in *Environmental and Resource Economics*

1 Introduction

This paper provides a decomposition analysis of the observed reductions of the concentrations of sulfur dioxide, nitrogen dioxide and ozone, in the twelve richest European countries. It quantifies the proportion of the reductions that can be attributed to fiscal policies, trade, and energy taxes. We show that fiscal spending policies and increasing trade openness explain the largest portion of the observed reductions of production-related pollutants (e.g., sulfur dioxide) while energy taxes explain most of the observed decreases of pollutants originated mostly in consumption activities (e.g., nitrogen dioxide). This is the first econometric study that compares the effects of fiscal expenditure policies, energy taxes and trade openness on environmental quality in Europe.

This analysis shows that a policy factor that so far has been largely neglected plays a key role in affecting pollution: fiscal spending. The impact of fiscal spending has proven to be important in most areas of the economy.[1] However the relationship between fiscal policies and environmental quality has received little attention in the literature.[2]

[1] Studies have focused for example on the effects of public expenditure level and composition on poverty reduction, income distribution and inequality (Kaplow, 2006), unemployment (Fougerè *et al.*, 2000), education (Hanushek, 2003), and many other areas.

[2] Exceptions are the theoretical models of Barman & Gupta (2010) and Gupta & Barman (2009) and López et al. (2011) which include a general equilibrium theoretical model and an empirical application.

The study of the impact of fiscal factors is likely to be especially important in Europe where the participation of government spending in the economy tends to be higher than in most other regions of the world (Dewan and Ettlinger, 2009).

In addition to fiscal spending, our analysis also examines how increases in trade intensity affect pollution in wealthy countries.[3] Earlier studies (such as Grossman and Krueger, 1992; Antweiler *et al.*, 2001; Frankel and Rose, 2005) have examined the effect of trade on pollution using samples that include a large proportion of middle income and poor countries. These studies have found that trade reduces pollution. The environmental improvements in middle and low income countries may be due to greater imports of cleaner technologies that already exist in rich countries (Antweiler *et al.*, 2001). However, it is possible that trade may not increase environmental efficiency in rich countries; it may merely induce them to shift production towards cleaner outputs thus displacing their dirty industries to poorer countries. Lastly, we analyze the effect of energy taxes and certain environmental regulations which may increase the incentives in rich countries to produce new and more environmentally efficient technology (Knigge and Görlach, 2005).

Empirical studies on trade and environment do not control for the level and composition of government spending and energy taxes. Hence, these studies may be affected by omitted variable bias as recognized by Antweiler *et al.* (2001). Typically two way fixed effects (TWFE) are used to deal with the bias; however this procedure is not efficient in controlling for country specific time-varying omitted variable bias.

[3] Gassebner et al. 2010 survey the literature and find that excluding OECD countries from the sample provides different results; specifically the relationship between GDP growth and pollution loses significance. This suggests that focusing on the richest countries of Europe could provide new insights about the relevance of fiscal policies, trade and energy taxes.

The literature analyzing the effects of energy and environmental taxes on pollution has mainly used simulation exercises, rather than econometric modeling (Baranzini *et al.*, 2000; Fullerton and Heutel, 2007, Fullerton et al., 2009). This is mainly due to the lack of suitable data that may capture the variability of institutions, regulations and enforcement variables that may affect pollution (Morley 2010). Another strand of literature has used firm or industry level data (Millock and Nauges, 2006; Morley, 2010). These studies find that energy and environmental taxes have a negative and significant impact on air pollutants. However, these studies may be affected by the issues concerning time-varying omitted variables..

Our study aims to empirically estimate the effects of the level and composition of government expenditures, trade, and energy taxes on three major air pollutants: sulfur dioxide (SO2), nitrogen dioxide (NO2) and ozone (O3). [4] SO2 is produced by industrial processes and electricity generation and is considered mainly a “production-generated pollutant.” A significant part of NO2 and O3 is produced by road vehicles implying that both pollutants may be considered as mainly (but not totally) “consumption-generated pollutants.”[5] We improve the analysis of pollution determinants regarding two other important aspects:

1. We introduce a method that generalizes the conventional Fixed Country Effects (FCE) approach; a method that we call Time-Varying Country-Specific Effects (TVCE). The TVCE method reduces the risk of spurious correlation between pollution and the explanatory variables of interest caused by time-varying as well as fixed unobserved or

[4] We select these pollutants because their measurements are reliable and consistent over time, they have the largest number of observations available, they can be regulated, and accepted quality standards exist for them (EPA, 2010).

[5] O3 is the product of the combination of Nitrogen Oxides (NOx) and volatile organic compounds

difficult-to-measure variables which may be correlated with the latter. While we directly control for certain environmental regulations, there may be other economic and institutional variables that may affect pollution which are either unobserved or difficult to measure, such as regulation enforcement that may change over time and is specific to each country. The TVCE is a parsimonious approach that allows for controlling for omitted variables without measuring them directly.[6]

2. We use a new dataset of air pollution for Europe. The existing empirical estimations have used the GEMS/AIR data which have observations for the period 1971-1996, (Grossman and Krueger, 1995; Antweiler *et al.*, 2001; Harbaugh *et al.*, 2002; Bernauer and Koubi, 2006). Our sample, using more recent data, has the advantage of including more monitoring stations in each of the countries analyzed, for the 1995-2008 period. The number of observations available for SO2 (about 16,000 observations distributed over 2,666 monitoring stations in 12 countries), for example, is five times larger than in the old data set. This large number of observations allows us to implement the TVCE method, which, as we shall see, requires us to estimate a large number of auxiliary coefficients.

The remainder of this paper is organized as follows: Section 2 discusses conceptual issues, Section 3 presents the econometric model, Section 4 describes the data, Section 5 summarizes the results and Section 6 concludes.

[6] An alternative method to control for time-varying unobservable variables is the so-called Added Controls Approach which sequentially introduces a large number of controls (Altonji *et al.*, 2005). Nevertheless, Altonji *et al.* (2005) do caution about this methodology: "….[it] is dangerous to infer too much about selection on the unobservables from selection on the observables if the observables are small in number and explanatory power or if they are unlikely to be representative of the full range of factors that determine an outcome". (p. 182).

2 Conceptual Issues

To analyze the impact of government spending composition, it is important to use a taxonomy of expenditures that is conceptually meaningful and consistent with the available data. López and Galinato (2007) proposed a taxonomy of government expenditures that distinguishes between expenditures on what they term "public goods," defined as those that alleviate the negative effects of market failures, and expenditures on "private goods," which do little to mitigate market imperfections.[7] Accordingly, government expenditures on public goods include expenditures on education, health, social transfers, environmental protection, research and development (R&D), knowledge creation and diffusion, as well as conventional public goods such as, institutions and law and order. By contrast, government expenditures on private goods are subsidies to special interest groups including credit and input subsidies, farm commodity programs, subsidies to the production and consumption of fossil fuels, industrial subsidies, and others.

Unlike government expenditures on private goods, expenditures on public goods may complement rather than substitute private sector spending. Household subsidies, both direct and indirect via education and health care provision, mitigate the negative effects of liquidity constraints on investments in human capital (e.g. Galor and Zeira, 1993) which according to recent studies affect a significant portion of households even in

[7] Fiorito and Kollintzas (2004) provide a different but related taxonomy based on the relationship between the types of goods provided by the government and private consumption. *Public goods* are defined as those that cannot be provided by the private sector such as defense, public order and justice. *Merit goods* include health, education and others that are in part provided by the private sector but where the public sector may have an important complementary role.

wealthy countries (Zeldes, 1989; Japelli, 1990; Grant, 2007; Attanasio *et al.*, 2008).[8] Investment in environmental protection, research and development, and creation and diffusion of knowledge, finance activities that otherwise would be under-funded due to generally insufficient market incentives for the private sector to invest in these areas (Dasgupta, 1996; Hoff and Stiglitz, 2000).

López et al. (2011) develop a theoretical model identifying the channels by which the level and composition of government spending may affect the environment. They find that the reallocation of government expenditure from private to public goods improves environmental quality indicators. They find that increasing the share of public goods expenditures on total government spending contributes to the expansion of aggregate output (or GDP), changes the composition of production towards human capital-intensive industries (and away from physical capital-intensive ones) and promotes investments in R&D. Hence, these factors induce three effects on production-generated pollutants: (1) Scale Effect: the expansion of aggregate output may increase pollution; (2) Composition Effect: a reduction in pollution due to the restructuring of production in favor of human capital–intensive activities that tend to pollute less than physical capital-intensive activities; (3) Technique Effect: increasing investments in R&D and in diffusion of knowledge, which may lead to the development of environmentally cleaner technologies.[9]

---

[8] In addition, studies have shown that human capital investments often have spillovers that increase their social value beyond their private returns (Blundell *et al.*, 1999; Fleisher *et al.*, 2010).

[9] Given that the output elasticities of energy (or electricity) range from 0.3 to 1.35 for OECD countries (Adeyemi and Hunt, 2007; Liu ,2004 and Olund, 2010), macroeconomic policies such as fiscal spending may have an impact on production related pollution by affecting its sources such as output from electric utilities and industry.

Furthermore, reallocating fiscal spending towards public goods may also reduce consumption-generated pollution by shifting consumption towards less polluting goods. For example, raising the share of public goods may entail greater investment in public transportation which substitutes for private transportation which, in turn, often implies less demand for energy and hence less pollution. In addition more investments in R&D increase the supply of fuel efficient cars and energy saving appliances including air conditioning and heating units.

The impact of trade expansion on the environment has also been associated with scale, technique and composition effects (Grossman and Krueger, 1992; Antweiler *et al.*, 2001; Frankel and Rose, 2005). The effects of trade vary depending on the nature of the pollutant and on the economy's level of income. Increases in the volumes of trade may cause an expansion of economic activity (scale effect) thus, ceteris paribus, raising production-generated pollution. Trade may also induce a technique effect on pollution but this effect has been mainly considered to be due to the fact that trade increases income, which in turn may raise the desire for stricter environmental regulations. Hence, if we control for real income, taxes and regulations, the effect of trade should capture mostly the output composition effect. Trade could also affect pollution by facilitating transfers of technology. The increased technology transfer effect is most important for poor countries that tend to be the ones that receive technologies from the more advanced countries. However, given that our sample includes only rich countries which are the ones that generate environmentally cleaner technologies, this effect should be limited in their own environments.

Environmental regulations and environmental taxes may have an effect on the environment mostly through the technique effect, by reducing the level of emissions per unit of goods produced or consumed (Knigge and Görlach, 2005). Environmental and energy taxes directly increase the costs of “dirty” inputs or of dirty consumption goods such as fuels or gasoline, thus inducing their savings and substitution. While these policies may also induce some output composition effect by increasing the relative price of outputs that use dirty inputs more intensively, this effect is likely to be weak. As Karp (2011) argues, one possible explanation for the weakness of the composition effect of environmental policies is that the costs of complying with environmental regulations account for only a small share of total production costs, creating little incentives to relocate production of dirty goods. Thus, unlike economy-wide policies, energy taxes and regulations are likely to have first order effects on techniques and the structure of consumption goods and only a second order effect on production composition.

Controlling for the scale effect (as we do in this paper), given the sample of rich countries that we use and the type of pollutants considered in our analysis, it is expected that energy taxes and environmental regulations mostly amplify the technique effect; trade mostly influences the composition effect in the case of production pollutants and has little effect on pollution produced by consumption activities. Fiscal policies may affect pollution via both the technique and output composition effects.

3 Econometric Model

We assume that the annual average pollutant concentration at monitoring station *i*, in country *j* at time *t*, Z_{ijt} , is determined by a vector reflecting the stocks of public and private goods provided by the government, $\mathbf{G_{jt}}$, trade intensity, TI_{jt}, country-specific energy taxes, M_{jt}, and environmental regulations at the country level, R_{jt}. In addition, we control for the three year moving average of per capita household final consumption expenditure (as a proxy for permanent per capita income), Y_{jt}. Additional controls include temperature (heating degree days), E_{jt}, and monitoring station characteristics, $\mathbf{X_{ij}}$. Finally, the model controls for unobserved monitoring station effects and time-varying unobserved country effects.

$$Z_{ijt} = \tilde{\psi}_{ij} + a_1\mathbf{G_{j,t-1}} + a_2 TI_{j,t} + a_3 M_{j,t} + a_4 R_{jt} + a_5 Y_{j,t} + a_6 E_{jt} + a_7 \mathbf{X_{ij}} + \zeta(\tau)_{jt} + \tilde{\varepsilon}_{ijt} \qquad (1)$$

$$i \in \{1, 2, ..., I\},\ j \in \{1, 2, ..., J\},\ t \in \{1995, ..., 2008\},$$

Where $\tilde{\psi}_{ij}$ is an unobserved monitoring station effect that can be fixed or random; $\zeta(\tau)_{jt}$ is a function of time that controls for fixed and time-varying country-specific effects; $\tau = t - 1995$; and $\tilde{\varepsilon}_{ijt}$ is an idiosyncratic error that is assumed to be independent and identically distributed with zero mean and fixed variance.

While we have data on government expenditure flows for various key components we do not have reliable measures of their respective stock levels, $\mathbf{G_{jt}}$. We thus write Equation (1) below in differences so that the annual differences of the government stocks can be approximated by the lagged level of corresponding government expenditure flows. We then have,

$$z_{ijt} = \psi_{ij} + \alpha_1 \mathbf{g}_{\mathbf{j,t-1}} + \alpha_2 ti_{j,t} + \alpha_3 m_{j,t} + \alpha_4 r_{jt} + \alpha_5 y_{j,t} + \alpha_6 E_{jt} + v_{jt} + \varepsilon_{ijt} \qquad (2)$$

where, $z_{ijt} \equiv Z_{ijt} - Z_{ij,t-1}$; $\mathbf{g}_{\mathbf{j,t-1}} \equiv \mathbf{G}_{\mathbf{jt}} \text{-} \mathbf{G}_{\mathbf{j,t-1}}$; $ti_{jt} \equiv TI_{jt} - TI_{j,t-1}$; $m_{jt} \equiv M_{jt} - M_{j,t-1}$; $r_{jt} \equiv R_{jt} - R_{j,t-1}$; $y_{j,t\text{-}1} \equiv Y_{jt} \text{-} Y_{jt\text{-}1}$; $v_{jt} \equiv \zeta(\tau)_{jt} - \zeta(\tau)_{jt-1}$; ψ_{ij} is an unobserved monitoring station effect; $\tilde{\varepsilon}_{ijt}$ is an idiosyncratic error that is assumed to be independent and identically distributed with zero mean and fixed variance.[10]

The v_{jt} effect corresponds to the TVCE. We approximate the v_{jt} effect by a (T-2)th order (country-specific) polynomial function of time,

$$v_{jt} = b_{0j} + b_{1j}\tau + b_{2j}\tau^2 + b_{3j}\tau^3 + \dots\dots + b_{T-2,j}\tau^{T-2} + \mu_{jt} \qquad (3)$$

where, $b_{0j}, b_{1j}, b_{2j}, \dots\dots, b_{T-2,j}$ are country-specific coefficients of the polynomial function of τ; μ_{jt} is the residual; T is the maximum number of observations for a country. Using (3) in (2) we obtain the estimating equation where the new disturbance term is $\varepsilon_{ijt} \equiv \tilde{\varepsilon}_{ijt} + \mu_{jt}$. The $(T-2)^{th}$ order (country specific) polynomial function of time in Equation (3) is the maximum order of approximation that allows for sufficient degrees of freedom to estimate the effects of observed country variables on pollution.

The TVCE method is related to methods used in the literature (Cornwell et al., 1990; Jacobsen et al., 1993; Friedberg, 1998; and Wolfers, 2006). However, these studies choose up to a second order polynomial of time to capture individual or region-specific slow moving omitted variables. The main advantage of the TVCE model proposed here is that it does not arbitrarily restrict the degree of approximation to a second order

[10] The fixed station characteristics $\mathbf{X}_{\mathbf{ij}}$ vanish as a consequence of first differencing.

polynomial as the earlier studies cited did. Instead the degrees of freedom in the data determine the limit of the time trend polynomial in the estimation. This approach allows for a much more flexible approximation of the omitted variables impact and hence reduces significantly the risks of omitted variable biases affecting the coefficients of the variables of interest.

There are a number of possible omitted variables that the function ν_{jt} may control for. The main omitted variables which we are concerned about are regulations and especially their degree of enforcement. There are several characteristics about regulations that do not lend themselves to be easily accounted for. They tend to suffer from constant revisions over time, for instance, existing regulations may include more sectors or new regulations may be proposed and adopted. The stringency of regulation enforcement might increase over time (but not necessarily linearly). Also, the evolution of enforcement of the same regulations may differ across countries. All these factors make regulation stringency and enforcement difficult to measure and therefore hard to control.

Regulations and their enforcement are specifically a concern for this study since they may follow similar patterns over time as the share of public expenditures over total government spending in many countries, which suggests a positive correlation between the omitted regulation and our variables of interest. This in turn implies that failing to control for the effect of the time-varying omitted variables may bias the coefficients of the relevant explanatory variables upwards (more negative).

Including our proposed TVCE approach, there are a few other candidate specifications for dealing with omitted variables bias. These specifications include (i) the standard country fixed effects model (FCE), (ii) country by year fixed effects which consists of fully interacted country and time dummies (iii) country specific trends which is essentially country dummies interacted with a time trend and (iv) the TVCE approach proposed above.

The limitation of the standard country fixed effects approach (i) is that it does not account for time varying omitted variables. Most of the potential omitted variables mentioned such as regulation stringency and enforcement are not fixed for each country and tend to vary over time. In contrast, our proposed TVCE approach, do control for time-varying omitted variables as long as the omitted variables exhibit some degree of systematic variations over time.

To fully control for the effects of the omitted variables it would be necessary to use the complete matrix of country-year dummies, which is approach (ii). There are two potential advantages of the TVCE approach with respect to using the country by year fixed effects model. Estimating a $(T-1)^{th}$ order polynomial function of time for each country is equivalent to using the complete matrix of country-year dummies because in this case we would estimate *T* independent parameters for each country (the country-specific constant term or fixed effect plus *T-1* parameters corresponding to the slopes of the polynomial), in total T x J independent parameters. Thus the TVCE specification that estimates a $(T-2)^{th}$ order polynomial has the advantage of estimating fewer parameters than approach (ii). The second and most important advantage of the TVCE approach over the country by year fixed effects is that in the latter it is impossible to estimate the effect of any observed countrywide variables as all variation in these explanatory variables is eliminated. For the TVCE approach we can use the $(T-2)^{th}$ order approximation which means that the TVCE approach comes close to a full country by year fixed effects model but also allows to estimate the effects of country-wide explanatory variables.[11]

Finally, the use of country specific time trends (iii) imposes a linear functional form of the omitted variables while the TVCE approach allows for a more flexible functional form using country specific polynomials of the time trend. It is also important to note that the TVCE approach is a generalization of both the standard FCE (i) and the country-specific time trends approach (iii)[12].

The TVCE method is indeed a generalization of the standard fixed-country effects model. It can be assumed that the FCE apply to estimations in levels and thus taking first differences, as in equation (2), would wipe them out. More generally, we may apply the FCE to a regression in differences instead of one merely in levels.[13] Applying FCE to first differences can be interpreted as a first order approximation of the

[11] The TVCE approach indeed follows the tradition of classical regression analysis of using prior information (or assumptions) as a means of economizing the number of parameters. For example, in pure cross-country regressions the full use of country effects would not allow estimating the effects of the (observed) variables of interest, and thus a common approach is to use regional effects instead of country effects. The prior information or assumption is that the countries within the region may have common unobserved effects.

[12] While our assumption that the unobserved time patterns can be fully captured by the (T-2)th polynomial approximation is not certain, we can test whether the μ_{jt} residuals (and therefore the $\varepsilon_{ijt} \equiv \tilde{\varepsilon}_{ijt} + \mu_{jt}$ error term) are time-independent. If the hypothesis that the residuals are time- independent is not rejected, then the $(T-2)^{th}$ order polynomial approximation may be sufficient to uncover the full time pattern of omitted variable effects on the endogenous variable. Hence, the TVCE approach would be effective in mitigating time-varying country-idiosyncratic biases caused by omitted variables. By contrast, rejection of the hypothesis of time-independent residuals would suggest that the effects of omitted variables are not fully controlled for.

[13] The inclusion of the FCE in regressions in differences has often been used in literature examining the determinants of economic growth (defined as log difference of per capita GDP), in which FCE are used to control for unobserved time-invariant country specific characteristics (see for example, Fölster and Henrekson, 2001 and Afonso and Furceri, 2010).

unobserved country effects.[14] In cases where the total number of time observations per country is greater than 3, the $(T-2)^{th}$ order approximation of the TVCE method is more general allowing for the FCE-in-differences estimators to be nested within the TVCE estimators. That is, the FCE-in-differences model can be tested as a special case of the TVCE by parametrically testing the following restrictions, $b_{1j} = b_{2j} = ... = b_{T-2,j} = 0$ for all $j \in \{1, 2, ..., J\}$, while $b_{oj} \neq 0$, for at least some j.

4 Data

The air pollution data consist of annual averages for SO2, NO2 and O3 observations, measured at a large number of monitoring stations in the 12 richest European countries for the 1995-2008 period. These air quality measures are taken from the AirBase dataset maintained by the European Environmental Agency. The complete list of countries is provided in Table D.1.

Government expenditure, household final consumption and trade data are obtained from the EUROSTAT database. We use the functional classification of government expenditures at the general government level.[15] The government expenditures on public goods include expenditures on public order and safety, environment protection, housing and community amenities, health, recreation, culture

---

[14] If the country effect in the level equation is $b_{00j} + b_{0j}\tau$ then by first differencing the regression, the level FE (b_{00j}) vanishes and the FCE applying to differences becomes b_{0j}.

[15] This classification, organizes government expenditures in ten general categories, according to their objectives or purposes, which is of the government's administrative or organizational structure. The categories are: (i) general public services, (ii) defense, (iii) public order and safety, (iv) economic affairs, (v) environmental protection, (vi) housing and community amenities, (vii) health, (viii) recreation, culture and religion, (ix) education and (x) social protection (Jacobs et al. 2009).

and religion, education and social protection. Trade intensity is defined as the sum of exports and imports of goods and services as proportion of GDP.

The implicit tax rate on energy is obtained from EUROSTAT Statistical Books (2010). The temperature indicator (heating-degree-days) is obtained from the EUROSTAT database[16]. Table D.2 presents the description and source of data, while Table D.3 provides summary statistics of the variables used in the regressions.

In the sample there is high variation in SO2 concentrations among monitoring stations and over time. Some stations have reported reductions in SO2 concentrations of up to 50% over the period. NO2 concentrations have not decreased as much as SO2. However NO2 concentrations have high variation across monitoring stations and countries. Ozone concentrations have increased in some stations located in Southern and Central Europe. See Table D.4 for measures of variability of the pollutants and Tables D.5, D.6 and D.7 for the annual averages of each pollutant across countries and years.

The main explanatory variables show a large degree of variation over time and across countries as shown in Figures D.1 and D.2. The share of public goods in total government expenditure has increased over time for most of the countries. Germany and Denmark have the highest shares reaching values of 0.78 to 0.79 in some years; while Belgium and the Netherlands exhibit the lowest shares of the period with values as low as 0.65. Similarly, the share of total government expenditure over GDP varies significantly across countries and within countries during the sample period. The countries with the lowest share of total expenditure over GDP are Spain and the United Kingdom, with

[16] We control for heating degree days in our estimation, but are unable to account for cooling degree days due to lack of data. However, this omission may not be so serious given that most countries in our sample have only small windows of time in the summer when air conditioners may be used, and even then their incidence is low. In contrast, heaters are used more intensively in the winters.

shares as low as 0.38 in some years, and the countries with the highest shares include Sweden, Finland and Denmark with shares as high as 0.62.

5 Estimation and Results

We estimate equation (2) after normalizing the total government expenditures by GDP and the government expenditures on public goods by total government expenditures. We also normalize trade intensity (exports plus imports) by GDP. These normalizations are convenient because they yield unit free measures of the variables, which diminish the problems of comparing currency values and inflation across time and countries.

We use a sixth order polynomial approximation for the time-varying country effects (equation 3). The reason for limiting the approximation to the sixth order is that in our unbalanced panel data there are countries for which we have only eight years of observations. This effectively implies that we can estimate a maximum of seven coefficients per country to capture the ν_{jt} effect (the b_{0j} and the six b_{tj} coefficients for each country) in order to preserve sufficient degrees of freedom to estimate the variables of interest.

The monitoring station effect ψ_{ij} may be uncorrelated with the observed explanatory variables in which case we can use a random station effects model. Alternatively, we may allow for arbitrary correlation between the unobserved monitoring station effect and the observed explanatory variables in which case we would need to use fixed monitoring station effects. We use both random station effects and fixed station effects in combination with time-varying country-specific effects (RSE-TVCE and FSE-TVCE, respectively). We present the results obtained using RSE-TVCE in Table 1 while Table 2 shows the FSE-TVCE estimators.[17] If the station effects are correlated with the explanatory variables, the FCE-TVCE would be consistent and the RSE-TVCE may not. However, the results from both estimators are statistically similar which means that both are consistent. Given that the RSE-TVCE estimators use both the within and between variation while FCE-TVCE only rely on within country variations, the former are more efficient (Kennedy, 2003; Wooldridge, 2002). We use the RSE-TVCE estimators for the subsequent analysis.

5.1 Specification Tests

5.1.1 Testing the Fixed Country Effect model.

We test the null hypothesis that $b_{1j} = b_{2j} = b_{3j} = b_{4j} = b_{5j} = b_{6j} = 0$ for all j which, as discussed earlier, is a test for the validity of the fixed country effects model. As indicated in the bottom of Table 1, the restricted model is rejected at the 1% level of significance in favor of the TVCE model for each of the three pollutants, meaning that the often used fixed country effect specification is statistically rejected.[18]

The coefficients $b_{1j}, b_{2j}, b_{3j}, b_{4j}, b_{5j}, b_{6j}$ are jointly significant at the 1% level of significance and the majority of them are individually significant. This, in conjunction

[17] The standard errors in all the estimates are robust to heteroskedasticity and autocorrelation.

[18] Additionally, we test whether the residuals from the RSE-TVCE estimations are time independent by regressing them on a time trend τ ($\varepsilon_{ijt} = constant + \beta\tau$).The null hypothesis that the residuals are time independent is not rejected at any reasonable level for any of the pollutants; p-values for the associated nulls are all above 0.99 (these results are available from the authors).

with the relatively large impact that including these effects has on the coefficient estimates of the key variables, reflects the importance of the RSE-TVCE approach.[19]

5.1.2 Reverse causality.

Consistent with the econometric model presented in Section 3, the normalized government expenditures are lagged in the model. This may avoid the direct reverse causality between these variables and the pollutant, often a source of biases in the estimated coefficients. In principle it would still be possible that such lagged expenditures are correlated with other concurrent omitted variables which would bias the coefficients. However, as we argued earlier, the country-specific time-varying effects largely minimize such a risk as these effects control for omitted variables.

It may be argued that reverse causality could be an issue for energy taxes as the tax variable is not lagged. Higher levels of pollution may be a factor that induces governments to raise energy taxes in which case there would be an upward (less negative) bias on the energy tax coefficient. However, it is unlikely that the level of energy taxes is influenced much by variations in local pollution as energy tax policies are mostly motivated in renewable energy and climate change policies rather than in local pollution-related objectives (Newberry, 2005; Biermann and Brohm, 2005; Decker and Wohar, 2007). But even if reverse causality were indeed an issue, the finding of a negative effect of the energy tax on pollution, as we do when using the RSE-TVCE estimates, would merely make such estimates a lower bound measure of the true effect and would not alter

[19] Table G.1 in Appendix G presents a summary of the analysis of the predicted values of the TVCE function. In most countries the effect of the omitted variables has been negative for SO2, and has changed sign over time for NO2 and O3. The majority of predicted values of the TVCE function are non-monotonic and have at least 2 turning points.

the sign of the estimates. That is, if we corrected for reverse causality bias, the estimates of the energy tax effect would be even more negative than the ones obtained in the estimations.

5.2 Analysis of the Estimates

5.2.1 Impact Analysis

The estimates indicate negative and significant effects of the government spending level and composition on SO2 and O3, and negative but not statistically significant effects on NO2, as shown by the coefficients of the share of expenditures in public goods over total government expenditure and the share of total government expenditure over GDP in Table 1. Trade shows a negative and significant effect on SO2 concentrations but not on the other pollutants, and energy taxes exert a negative and significant effect on NO2 but not on the other contaminants.

In Table 3 we show the elasticities for the main determinants of each pollutant. The importance of each of these effects is also expressed by the relative changes within the sample (impact of changing the explanatory variables in one standard deviation, expressing it as proportion of the sample standard deviations of the pollutant). Increasing the share of government expenditures on public goods by 1%, holding total government expenditure constant, may result in a 3.9% reduction of SO2 concentrations and a 1.25% decrease in O3 concentrations. Increasing the share of expenditures on public goods by one standard deviation reduces SO2 concentrations by 22.7% and O3 by 19.4% of their respective standard deviations.

The concentrations of SO2 and O3 may be reduced by 2.6% and 0.82% respectively if total government expenditure increases by 1%. The increase of one standard deviation of the share of total expenditure with respect to GDP may result in a standard deviation reduction of 35.1% for SO2 and 31.7% in the case of O3.

We find that the elasticity of energy taxes is -0.31 for NO2. If the energy tax rates increase by one standard deviation, the concentration of NO2 may be reduced by 12% of its standard deviation.[20] The estimated effects of energy taxes are not significant for SO2 which is caused mainly by industrial processes and electricity generation. We did not find statistically significant effects of energy taxes on O3 which is formed by certain precursor gases in combination with weather conditions.[21] Our results suggest that energy taxes only affect pollution levels caused mainly by road and off-road fuel consumption.

Trade has a negative and significant effect on SO2 but does not have a significant effect on neither NO2 nor O3 concentrations. The estimates imply that increasing trade intensity by 1% may result in a 1.1% reduction of SO2 concentrations. If trade intensity is increased by one standard deviation, SO2 concentrations are reduced by 49% of its standard deviation. Hence, as predicted by our conceptual analysis, trade affects

---

[20] These findings are consistent with the elasticity estimates in a few studies that have measured these effects. Millock and Nauges (2006) estimate elasticities of energy taxes on NO2 and SO2 that vary from -2.7 to -0.2 depending on the industry analyzed.

[21] One of the reasons energy taxes do not have a significant effect on O3 concentrations might be the nature of this pollutant since it is not emitted directly by any source and it is rather formed by the combination of certain precursor gases especially under hot and sunny weather conditions (EEA, 2007). Another possible reason might be the positive effect of energy taxes over the participation of diesel vehicles on the automobile fleet (data that is not available for all countries and time periods); diesel vehicles tend to emit three times more ozone-precursor gases than gasoline vehicles. Vestreng *et al.*, 2008 has shown that this is true in the countries where systematic data on the participation of diesel vehicles are available.

"production" pollutants most likely through the composition effect but does not affect "consumption" pollutants.

The coefficients for the level of per capita household consumption are positive in our estimates and mostly significant while most existing empirical studies for high income countries obtain a negative effect on local pollutants (Antweiler et al. 2001; Bernauer and Koubi 2006; Deacon and Norman, 2007). This divergence may stem from our effort to mitigate the omitted variable biases by controlling for energy taxes, environmental regulation and other unobserved economy-wide variables that may be positively correlated with per capita income or consumption and that have a negative impact on pollution. The standard estimates are likely to attribute the effects of these variables to per capita income and thus conclude that increasing per capita income or household consumption may reduce pollution. By contrast our estimates isolate the pure effect of income or consumption on pollution.

5.2.2 Comparison across different specifications

In the conceptual section we justified the use of the TVCE model and indicated its advantages over alternative specifications. Tables E.1, E.2 and E.3 present the results of RSE, RSE with Fixed Country Effects (RSE-FCE), and RSE-TVCE with the full range of orders of approximation from 1 to 6.[22] The sign and significance of the results of our main coefficients are retained as they stabilize with higher polynomials of the time trend. A few results are however worth noting. As we control for time varying omitted variables and increase the order of approximation, both the magnitude and significance of the share

[22] The results for the Fixed Monitoring Stations Effects with Time Varying Country Effects (FSE-TVCE) alongside alternative specifications are available from the authors.

of government expenditures on public goods increase. This may imply that the higher order approximations of the TVCE are capturing unobserved country level variables that are correlated with spending and pollution . Similarly, we find that in our base estimations trade has no significant effect on NO2 and O3 pollutants. However, using simple RSE-FCE estimations the effect of trade becomes negative and significant. This may imply that our TVCE approach may be capturing country-wide omitted variables such as the degree of regulations enforcement which may be biasing upward the RSE-FCE estimations of the impact of trade on NO2 and O3.

5.2.3 Sensitivity Analysis

In addition to the specification tests reported earlier, we performed a series of sensitivity analyses to ascertain the robustness of the estimators. We checked for extreme data points that may dominate the sign and significance of key estimates and looked for individual country dominance.

We conducted two types of dominance tests. In order to account for extreme data points, we first re-estimated the model by excluding observations in the top and bottom 1% of the share of government expenditures on public goods. The same procedure is followed by re-estimating the model without observations in the top and bottom 1% of the energy tax rate, pollutant concentration and trade intensity. The parameters are robust to the sample changes, except for the case of trade when dropping the bottom 1% of SO2 observations. This result indicates that the effect of trade is weak even on production pollutants, once we control for energy taxes, environmental regulations, fiscal expenditure as well as other unobserved factors. Signs, significance and magnitudes of the parameter estimates from these models are shown in Tables F.1 to F.6 in Appendix F.

The second type of tests focuses on the effect of potential country dominance. We re-estimated our benchmark models, dropping one country at a time, to check whether they alter the parameter estimates of the share of public goods (for the SO2 and O3 regressions), of the energy taxes (for the NO2 regressions) and of trade (for SO2). As shown in Figures F.1 to F.4 in Appendix F, removing one country at a time does not affect the sign and significance of the estimated parameters, with the exception of share of public goods over government expenditure in the O3 regression, which seems to be dependent on Italy.[23]

5.3 Decomposition Analysis.

Table 4 shows the average annual changes in all pollutants for the analyzed period and the decomposition of fiscal, trade and environmental policy effects on each one of them. SO2, mainly a production pollutant, has decreased very rapidly over the period at an annual rate of 8.5% but NO2 and O3, considered mainly consumption pollutants, have not improved nearly as much. NO2 concentrations have fallen by only 1.4% per annum and O3 concentrations have increased in almost all countries showing an average annual rate of increase of 0.9%.

As mentioned in the conceptual section, we expect fiscal policies to mostly affect air pollution concentration via the composition and technique effects, with trade having a

[23] It is worth noting that Italy includes a large number of observations, more than 1,500 observations or about 8% of the total. This does not necessarily indicate a lack of robustness of the estimators; it is indeed remarkable that the coefficients are robust to the exclusion of all other countries even if dropping individual countries often entails removing 7% or more of the total observations.

larger effect through the former on production pollutants and a negligible effect through the latter on consumption pollutants. We also expect that environmental policies and energy taxes would affect consumption pollution mainly via the technique effect. As can be seen in Table 4, these predictions are fully corroborated by the empirical results.

SO2 reductions are mostly explained by trade and fiscal policies, which together explain practically all the observed reductions, meaning that without those policies SO2 levels would have increased over the analyzed period. The large contribution of trade and of the increased share of expenditures in public goods over total government expenditure and of the share of total government expenditure over GDP may be the result of the production shift towards cleaner, possibly more human capital-intensive industries. Environmental regulation, specifically the "Large Combustion Plant Directive" (see Table A.2 for a detailed description of this regulation), also contributes to the reduction in SO2 concentrations possibly through a technique effect.

In the case of NO2, energy taxes explain a major part of the observed modest reduction; about 52% of this reduction is most likely due to their direct effect causing higher energy prices and hence less consumer demand for energy. They may also induce a technique effect that reduces pollution. This suggests that energy taxes are an effective instrument to reduce this type of pollutant, and reflects the European countries' demand (on average) for less NO2 emissions per unit of goods consumed.

Fiscal policies associated with an increased participation of government spending in GDP and progressive shifts towards the provision of public goods have a strong (unintended) effect towards reducing ozone. In fact, the combined effect of the observed fiscal spending policies in Europe has been to induce a reduction of ozone concentrations by more than 1% per annum. That is, if Europe had not increased the share of government spending in GDP and if it had not changed the spending composition towards public goods, ozone would have increased twice as fast as what was actually experienced. The European fiscal spending policies may explain why in these countries O3 concentrations have not increased nearly as much as in other regions of the world. Additionally, fiscal policies are the only policies considered that have any effect on ozone, which is probably the most difficult to control among the measured pollutants.

6 Conclusion

This study finds that fiscal, trade and energy tax policies implemented by the twelve richest European countries are important determinants of pollution through various mechanisms. Large and increasing public sector participation and increasing prioritization of public goods over private goods in the European countries analyzed have had a hitherto ignored effect by reducing the concentrations of sulfur dioxide and ozone but not nitrogen dioxide. In addition, we find that the high energy tax policy adopted by the majority of the European countries over the last few decades have substantially contributed to reduce the concentrations of nitrogen dioxide but have no effect on ozone and sulfur dioxide. Finally, trade openness has a direct effect on sulfur dioxide but no effect on nitrogen dioxide or ozone.

These results should be regarded as an added incentive for EU countries to at least persist if not increase the emphasis on fiscal policies and energy taxes that trigger the development of new technologies. The study may also present an argument for other countries which have not yet adopted these policies to implement them. The results have implications for several non-European countries including the USA and large developing countries which currently have much lower energy taxes and fiscal spending policies that are heavily oriented to provide private goods instead of public goods. Pursuing fiscal policies as adopted by some European countries may potentially have a large unintended environmental pay-off.

To the best of our knowledge this is the first paper that systematically examines the role of fiscal spending policy, trade and energy taxes on Europe’s environmental quality, using a methodology that obtains estimates mostly free from time-varying omitted variable biases.

REFERENCES

Adeyemi, O. I., and L.C. Hunt (2007), "Modelling OECD Industrial Energy Demand: Asymmetric Price Responses and Energy-Saving Technical Change," Energy Economics 29(2007): 693-709

Afonso A, Furceri D (2010) Government size, composition, volatility and economic growth. Europ. J. Polit. Economy 26:517-532

Antweiler W, Copeland B R, Taylor M S (2001) Is Free Trade Good for the Environment? Amer. Econ. Rev. 91:877-908

Altonji J G, Elder T E, Taber C R (2005) Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. J. Polit. Economy 113:151-184

Attanasio O P, Goldberg P K, Kyriazidou E (2008) Credit Constraints in the Market for Consumer Durables: Evidence from Micro Data on Car Loans. Int. Econ. Rev. 49:401-436

Baranzini A, Goldemberg J, Speck S (2000) A future for carbon taxes. Ecolog. Econ. 32:395-412

Barman T.R., Gupta M.R (2010) Public Expenditure, Environment, and Economic Growth, Journal of Public Economic Theory 12:1109-1134

Bernauer T, Koubi V (2006) States as Providers of Public Goods: How Does Government Size Affect Environmental Quality? Working Paper No. 14, Center for Comparative and International Studies (ETH Zurich and University of Zurich). http://ssrn.com/abstract=900487

Biermann F, Brohm R (2005). Implementing the Kyoto Protocol without the United States: The Strategic Role of Energy Tax. Adjustments at the Border. Climate Policy 4:289–302

Blundell R, Dearden L, Meghir C, Sianesi B (1999) Human Capital Investment: The Returns from Education and Training to the Individual, the Firm, and the Economy. Fisc. Stud. 20:1-23

Cornwell, C., P. Schmidt, and R. C. Sickles (1990), "Production Frontiers With Cross-Sectional and Time-Series Variation in Efficiency Levels." Journal of Econometrics 46 (1990): 185-200.

Dasgupta P (1996) The Economics of the Environment. Environ. Devel. Econ. 1:387–428

Deacon, R. T. and Norman, C. (2007) “Is the environmental Kuznets curve an empirical regularity?” In R. Halvorsen and D. Layton, *Frontiers of Environmental and Natural Resource Economics*, Edward Elgar Publishers, U.K.

Decker C S, Wohar M E (2007) Determinants of state diesel fuel excise tax rates: the political economy of fuel taxation in the United States. Ann. Reg. Sci. 41:171–188

Dewan S, Ettlinger M (2009) Comparing Public Spending and Priorities Across OECD Countries. Center for American Progress. www.americanprogress.org

European Environment Agency (2007) Air pollution in Europe 1990–2004. EEA Report No. 2/2007

Environmental Protection Agency (EPA) (2010) Available at: http://www.epa.gov/air/urbanair/

EUROSTAT Statistical Books (2010) Taxation Trends in the European Union Data for the EU Member States and Norway

Fiorito R, Tryphon K (2004) Public goods, merit goods, and the relation between private and government consumption. Europ. Econ. Rev. 48:1367-1398

Fleisher B, Li H, Zhao M Q (2010) Human Capital, Economic Growth, and Regional Inequality in China. J. Devel. Econ. 92:215-231

Fölster S, Henrekson M (2001). Growth effects of government expenditure and taxation in rich countries. Europ. Econ. Rev. 45: 1501-1520

Fougère D, Kramarz F, Magnac T (2000) Youth employment policies in France. Europ. Econ. Rev. 44, 928-942

Frankel J A, Rose A K (2005) Is Trade Good or Bad for the Environment: Sorting Out the Causality. Rev. Econ. Statist. 87:85–91

Friedberg, L. (1998), "Did Unilateral Divorce Raise Divorce Rates? Evidence from Panel Data." American Economic Review 88(3): 608-627.

Fullerton D, Heutel G (2007). The General Equilibrium Incidence of Environmental Taxes. J. Public Econ. 91:571-591

Fullerton D, Leicester A, Smith S (2009) Environmental Taxes. In: Institute for Fiscal Studies (IFS) (Eds.), Dimensions of Tax Design. Oxford: Oxford University Press, pp428-521 http://works.bepress.com/don_fullerton/37

Galor O, Zeira J (1993) Income Distribution and Macroeconomics. Rev. Econ. Stud.:60, 35–52

Gassebner M, Lamla M, Sturm J E (2010) Determinants of pollution: what do we really know? Oxf. Econ. Pap. (in press). DOI 10.1093/oep/gpq029

Grant C (2007) Estimating credit constraints among US households. Oxf. Econ. Pap. 59:583-605

Grossman G M, Krueger A B (1992) Environmental impacts of a North American Free Trade Agreement. Centre for Economic Policy Research Discussion Paper 644

Grossman G M, Krueger A B (1995) Economic Growth and the Environment. Quart. J. Econ. 112:353-378

Gupta M.R, Barman T.R., (2009) Fiscal Policies, Environmental Pollution and Economic Growth. Economic Modelling 26:1018-1028

Hanushek E (2003) The Failure of Input Based Schooling Policies. Econ. J. 113(485):F64-F98

Harbaugh W, Levinson A M, Wilson D M (2002) Reexamining the empirical evidence for an Environmenal Kuznets Curve. Rev. Econ. Statist. 84:541–551

Hoff K, Stiglitz J (2000) Modern Economic Theory and Development. In Meier, G., Stiglitz J (Eds.), Frontiers of Development Economics. New York: Oxford University Press and the World Bank, pp389-459**.**

Jacobs D, Helis J, Bouley D, (2009) Budget Classification, Technical Notes and Manuals. Washington: International Monetary Fund.

Jacobsen L. S., R. J. Lalonde, and D. G. Sullivan (1993), "Earnings Losses of Displaced Workers." American Economic Review 83(4): 685-709.

Jappelli T (1990) Who is Credit Constrained in the US Economy? Quart. J. Econ. 105:219-234

Kaplow L (2006) Public goods and the distribution of income. Europ. Econ. Rev. 50:1627-1660

Karp L (2011) The Environment and Trade: a Review. Department of Agricultural and Resource Economics, University of California, Berkeley, http://areweb.berkeley.edu/~karp/environmenttradejan24.pdf

Kennedy, P., A (2003) Guide to Econometrics, MIT Press, 5th Edition.

Knigge M, Görlach B (2005) Effects of Germany’s Ecological Tax Reforms on the Environment, Employment and Technological Innovation: Summary of the Final Report of the Project. Ecologic Institute for International and European Environmental Policy, Berlin, August 2005

Liu, G. (2004), "Estimating Energy Demand Elasticities for OECD Countries: A Dynamic Panel Data Approach," Statistics Norway Research Department, Discussion Paper No. 373.

López R, Galinato G (2007) Should Governments Stop Subsidies to Private Goods? Evidence from Rural Latin America. J. Public Econ. 91:1071-1094

López R, Galinato G, Islam A (2011) Fiscal Spending and the Environment: Theory and Empirics. Journal of Environmental Economics and Management 62(2):80-198

Millock K, Nauges C (2006). Ex-Post evaluation of an earmarked tax on air pollution. Land Econ. 82:68-84

Morley B (2010) Empirical Evidence on the Effectiveness of Environmental Taxes. Working Paper No. 02/10. Department of Economics, University of Bath, Bath, UK

Newberry D (2005) Why Tax Energy? Towards a More Rational Policy. Energy J. 26:1-40

Olund, K. (2010) "The Industrial Electricity Use in the OECD Countries," Lulea University of Technology thesis, 2010

Vestreng V, Ntziachristos L, Semb A, Reis S, Isaksen I S A, Tarrasón L (2008) Evolution of NO_X emissions in Europe with focus on road transport control measures. Atmospheric Chemistry and Physics Discussions 8:10697-10747

Wolfers, J. (2006), "Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results." American Economic Review 96(5): 1802-1820.

Wooldridge, J. M. (2002), Econometric Analysis of Cross Section and Panel Data, Cambridge,

Zeldes S P (1989) Consumption and Liquidity Constraints: An Empirical Analysis. J. Polit. Economy 97:305-346

Table 1

Random Monitoring Stations Effects with Time Varying Country Effects (RSE-TVCE)

	Ln Diff SO2	Ln Diff NO2	Ln Diff O3
Share of expenditures in public goods over total government expenditures (lagged)	-5.33** [1.27]	-0.19 [0.49]	-1.69** [0.52]
Share of total government expenditures over GDP (lagged)	-5.52** [1.78]	-0.37 [0.71]	-1.73* [0.77]
Time difference of Energy Tax Rate	-0.11 [0.06]	-0.18** [0.03]	0.057 [0.04]
Time difference of Regulation over large Plants	-0.49** [0.07]		
Time difference of Regulation over NOx		-0.34 [0.51]	1.54 [0.83]
Time difference of Log of Trade (X+M)/GDP	-1.13** [0.40]	-0.41 [0.25]	-0.21 [0.43]
Time difference of 3-Year Moving Average of Ln of Household final consumption per capita	0.053** [0.02]	0.004 [0.01]	0.026* [0.01]
Number of Observations	16,222	19,374	15,282
No. of Monitoring Stations	2,666	3,176	2,274
Overall R-Squared	0.11	0.06	0.10
Specification tests: Testing the fixed country effects-random site effects model: Log Likelihood Ratio Test Ho: $b_{1j} = b_{2j} = ... = b_{T-2,j} = 0$	426**	322**	316**

Robust standard errors in brackets.

* significant at 5%; ** significant at 1%

Not reported in the table are 77 coefficients for each equation for the variables that capture the TVCE, 12 coefficients for year effects and one coefficient for heating degree days.

Table 2
Fixed Monitoring Stations Effects with Time Varying Country Effects (FSE-TVCE) Regressions

	Ln Diff SO2	Ln Diff NO2	Ln Diff O3
Share of expenditures in public goods over total government exp (lagged)	-5.92** [1.34]	-0.27 [0.54]	-1.68** [0.58]
Share of total government expenditure over GDP (lagged)	-6.17** [1.91]	-0.12 [0.79]	-1.61 [0.86]
Time difference of Energy Tax Rate	-0.14* [0.07]	-0.20** [0.034]	0.05 [0.04]
Time difference of Regulation over large Plants	-0.47** [0.08]		
Time difference of Regulation over NOx		-0.55 [0.66]	1.70 [1.06]
Time difference of Log of Trade (X+M)/GDP	-1.09* [0.48]	-0.37 [0.27]	-0.23 [0.48]
Time difference of 3 Year Moving Average of Ln of Household final consumption per capita	0.06* [0.02]	0.04** [0.01]	0.05** [0.01]
Number of Observations	16,222	19,374	15,282
No. of Monitoring Stations	2,666	3,176	2,274
Overall R-Squared	0.12	0.07	0.13
Specification tests: Testing the fixed country effects-random site effects model: Log Likelihood Ratio Test Ho: $b_{1j} = b_{2j} = ... = b_{T-2,j} = 0$	463**	361**	403**

Robust standard errors in brackets.
* significant at 5%; ** significant at 1%
Not reported in the table are 66 coefficients for each equation for the variables that capture the TVCE, 12 coefficients for year effects and one coefficient for heating degree days.

Table 3
Elasticities and Sample Quantitative Effects

	SO2	NO2	O3
Elasticity of the Share of Public Goods	-3.91**	n. s.	-1.25**
Change in the pollutant when the Share of Public Goods increases by one Standard Deviation (% of std dev of pollutant)	-22.70%**	n. s.	-19.45%**
Elasticity of the ratio of total government expenditure over GDP	-2.63**	n. s.	-0.82*
Change in the pollutant when the ratio of total government expenditure over GDP increases by one Standard Deviation (% of std dev of pollutant)	-35.10%**	n. s.	-31.37%*
Elasticity of the Energy Tax Rate	n. s.	-0.31**	n. s.
Change in the pollutant when the Energy Tax Rate increases by one Standard Deviation (% of std dev of pollutant)	n. s.	-12.32%**	n. s.
Elasticity of Trade	-1.13**	n. s.	n. s.
Change in the pollutant when Trade increases by one Standard Deviation (% of std dev of pollutant)	-49.23%**	n. s.	n. s.

* significant at 5%; ** significant at 1%
n. s.: non-significant

Table 4
Decomposition analysis of the effect of the various factors

	Observed annual average rate of growth of the pollutant (%)	Annual average contribution (in percentage points)				
		Fiscal Policy		Environmental Policy		Trade Policy
		Share of government provided public goods over total expenditures	Share of total government expenditure over GDP	Regulation	Energy taxes	
SO2	-8.51	-5.56*	-2.27*	-3.79*	n. s.	-2.76*
NO2	-1.37	n. s.	n. s.	n. s.	-0.58*	n.s.
O3	0.91	-0.39*	-0.71*	n. s.	n. s.	n.s.

Note: The rates of growth used to create this table were calculated as the annual average growth. In the case of the pollutant the annual rate of growth of each monitoring station was calculated, and then a country average was taken for each country and finally the average over all of the years available in the sample. For the rest of the variables at the country level, first the rate of growth with respect to the previous year was calculated then the average of the whole period.
*Significant to at least 5%

Chapter 3: Caught in a Productivity Trap: A Distributional Perspective on Gender Differences in Malawian Agriculture

By: Amparo Palacios López[1]

"[C]hildren shrieking at play; and women bent double - most with infants slung on their backs - hoeing the corn and beans; and the men sitting in the shade stupefying themselves on chibuku, the local beer, or kachasu, the local gin."

Paul Theroux, Dark Star Safari: Overland from Cairo to Cape Town (2002)

"While a great deal has been learned about what works and what does not when it comes to promoting greater gender equality, the truth remains that progress is often held back by the lack of data or adequate solutions to the most 'sticky' problems."

The World Bank World Development Report 2012 *Gender Equality and Development*

1 Introduction

Globally, 1.4 billion people, or one quarter of the population of the developing world, live in extreme poverty, and an additional 1.2 billion live in moderate poverty. The analysis of regional contributions to global poverty indicates that although sub-Saharan Africa represents only 12 percent of the world population, it accounts for 27 percent of the global poor, and that poverty in sub-Saharan Africa is being reduced at a much slower pace than elsewhere (Chen and Ravallion, 2008).[2] Aggregate agricultural growth has

[1] With Talip Kilic and Markus Goldstein

[2] The poverty rate in sub-Saharan Africa is estimated to have declined only 3 percentage points, from 54 to 51 percent, throughout the period of 1985-2005.

been documented to bring disproportionate gains to the poorest in the developing world.[3] In sub-Saharan Africa, nearly 75 percent of the extreme poor reside in rural areas, and 91 percent of the rural extreme poor are estimated to participate in agriculture. As smallholder agriculture is the predominant form of farm organization in the region (FAO, 2009), smallholder agricultural productivity growth has been identified as a key driver of poverty reduction and increased food security.[4] In targeting sustainable poverty gains through smallholder-based agricultural growth, national development plans across sub-Saharan Africa have emphasized the reduction of gender differences in agricultural productivity. Most recently, FAO (2011) asserted that if female farmers had the same access to productive resources as men, they could increase yields by 20 to 30 percent, which could increase total agricultural output in developing countries by 2.5 to 4 percent and lift 100 to 150 million people out of hunger. Increased productivity among female farmers is also often argued to result in double-barreled payoff: (i) poverty alleviation through positive impact on overall smallholder productivity growth, and (ii) improved development outcomes for the next generation.[5]

Although the estimates of gender differences in agricultural productivity (henceforth referred to as the gender gap)[6] across sub-Saharan Africa range widely from 4 to 40 percent, the majority cluster around 20 to 30 percent. The studies that compare

[3] Ligon and Sadoulet (2008) document that a 1 percent rise in agricultural GDP results in 6 percent income growth for the lowest income decile of the population.
[4] Irz et al. (2001) estimate that for every 10 percent increase in farm yields, there has been a 7 percent reduction in poverty in sub-Saharan Africa.
[5] See WB (2011), and Doepke and Tertilt (2011) for a review.
[6] Agricultural productivity is commonly proxied by major crop production quantity per hectare or gross value of crop output/profit per hectare.

productivity outcomes on female- vs. male-managed plots across and within households provide further support for the presence of systematic and persistent gender differences in agricultural productivity in the region (Akresh, 2005; Alene et al., 2008; Gilbert et al., 2002; Goldstein and Udry, 2008; Moock 1976; Peterman et al., 2011; Oladeebo and Fajuyigbe, 2007; Quisumbing et al., 2001; Saito et al., 1994; Tiruneh et al., 2001; Udry, 1996; Vargas Hill and Vigneri, 2011). The major reasons for the observed gender gap have been identified as gender differences in (i) access to and use of agricultural inputs, (ii) tenure security and related investments in land and improved technologies, (iii) market and credit access, (iv) human and physical capital, and (v) informal institutional constraints affecting farm/plot management and marketing of agricultural produce.[7] Regardless of whether the comparisons are made across or within households, the common thread across the relevant literature is that the gender gap disappears or diminishes significantly once the researcher controls for the factors discussed above.

Despite what could be perceived as a well-established evidence base on the extent and proximate causes of the gender gap across sub-Saharan Africa, the overwhelming majority of empirical studies on the topic have used data from small-scale surveys that were limited in terms of geographic coverage, topic, or attention to intra-household dynamics (or, in some cases, all three). With the exception of Akresh (2005), none of the above-referenced papers rely on nationally-representative survey data. Dearth of nationally-representative, methodologically-sound data collected in heterogeneous

[7] Cultural roles that are assigned to males and females regarding domestic duties and those that may underlie the gender segregation in crop production (i.e. staple vs. cash crop cultivation, high-yielding vs. low-yielding variety cultivation, etc.) could be thought of as informal institutional constraints.

settings across sub-Saharan Africa has in turn inhibited the computation of externally-valid, rigorous estimates. Our study seeks to start filling this gap by providing a nationally-representative analysis of the gender gap in Malawi, using a different econometric approach than existing studies.

Our econometric approach is underlined by the use of an identification strategy that has been utilized extensively in labor economics since the seminal studies of Oaxaca (1973) and Blinder (1973), most notably in the analyses of the gender wage gap, union wage gap, and growing wage inequality. Specifically, we decompose the average difference in agricultural productivity between male-managed and female-managed plots into (i) the portion that is driven by gender differences in levels of observable attributes (i.e. *the endowment effect*), and (ii) the portion that is driven by gender differences in returns to the same set of observables (i.e. *the structure effect*). To our knowledge, this is the first time this method has been applied to understanding the gender gap.

Complementing this aggregate decomposition analysis, we provide a detailed decomposition of the mean gender gap, identifying the contribution of each observable covariate towards the endowment and structure effects. In contrast with the available microeconomic evidence, the detailed decomposition documents, within a partial-equilibrium framework, the relative quantitative importance of each factor in explaining the mean gender differential. This in turn facilitates further analysis to identify the causes of differences in key factors contributing to the gender gap so that the emerging insights could inform the design of policy interventions addressing the gender gap at its roots.[8]

The second contribution of our study relates to the application of the decomposition methodology to distributional statistics beyond the mean through the use of recentered influence function (RIF) regressions. Since key contributors towards the gender gap might differ across farmer subpopulations of varying productivity levels, the RIF decomposition is a useful tool for tracing out the heterogeneity in constraints faced by farmers with different gender and productivity profiles, and thus, tailoring better targeted policies that are underlined by analyses that move beyond the "average" male vs. female farmer. Towards this end, we carry out (i) the aggregate decomposition of the gender gap at each decile of the agricultural productivity distribution, and (ii) the detailed decomposition of the gender gap at the 10^{th}, 50^{th}, and 90^{th} percentiles. The paper also discusses the changes in the shares of endowment and structure effects as part of the aggregate decomposition, and the variations in the contributions of key factors towards the endowment and structure effects at selected percentiles.

Finally, the multi-topic and national-representative nature of our household survey data represents the third contribution to the literature on the gender gap in sub-Saharan Africa. The availability of geo-referenced household and agricultural plot locations also allows us to create synergies with geographic information system (GIS)

[8] For instance, if the researcher confirms that female managers, on average, have access to less inorganic fertilizer, and that the gender differences in inorganic fertilizer application is a key contributor towards the gender gap, it becomes crucial to understand why female managers have access to less inorganic fertilizer so that the policy interventions could target the underlying causes of this phenomenon.

data for the purpose of incorporating relevant geospatial variables into the modelling efforts.

There are five key findings from our study. First, on average, female-managed plots in Malawi are 25 percent less productive than those that are managed by males. Second, 82 percent of the mean gender gap is explained by the differences in observable covariates, i.e. the endowment effect. The direct pay-off to addressing market and institutional failures that affect men and women differentially is economically significant: ensuring that female plot managers have similar years of schooling and apply similar levels of non-labor agricultural inputs, including inorganic fertilizer, pesticides/herbicides, and improved and/or export crop varieties could reduce the mean gender gap by 50 percent. Deficiencies on female-managed plots regarding household adult male labor input and access to agricultural implements are other key factors exacerbating the gender gap. Third, the remaining 18 percent of the mean gender gap is mostly explained by gender differences in *returns* to (i) household adult male labor input and inorganic fertilizer application, which have significantly lower positive effects on the productivity of female-managed plots, and (ii) the child dependency ratio, which has a highly significant and negative effect on the productivity of female-managed plots, in contrast to no effect on the productivity of male-managed plots. Fourth, the gender gap increases significantly across the agricultural productivity distribution: the differential stands at 22 and 37 percent at the 10^{th} and 90^{th} percentile, respectively. Finally, we find that the gender gap is explained predominantly by the endowment effect in the first half of the agricultural productivity distribution, with the endowment effect still explaining close to 90 percent of the gender gap at the median. Above the median, however, the contribution of the endowment effect towards the gender gap declines steadily such that the structure effect culminates in explaining 34 percent of the gender gap at the 90th percentile.

The rest of the paper is organized as follows. Section 2 presents a review of the evidence on the gender gap in sub-Saharan Africa. Section 3 provides an overview of the Malawian context, and describes the data. Sections 4 and 5 present the mean decomposition methodology and the results from the mean decomposition, respectively. Likewise, Sections 6 and 7 present the RIF decomposition methodology and the results from the RIF decomposition, respectively. Section 8 offers concluding remarks and expands on the policy implications of our findings.

2 Gender Differences in Agricultural Productivity in Sub-Saharan Africa : Review of Evidence

The studies that investigate the gender gap in sub-Saharan Africa are quite heterogeneous in terms of the type of data and the estimation strategies that they use. The existing literature broadly features two strands. The first strand is composed of studies that conduct their analyses at the household-level and do not link plot-level outcomes to the identity of the managers and/or owners within study households. The second strand is composed of a handful of empirical studies that use plot-level data linked to individual managers within study households. Across these strands, the relevance and applicability of the results for policy have been limited due to shortcomings in terms of questionnaire design, empirical methodology, and/or sample representativeness. “[T]he inconclusiveness of gender research due to either methodological or data limitations obscures the policy and programmatic recommendations that emerge from gender productivity analysis, and do not enable us to ascertain whether gender matters in producing evidence-based agricultural policy” (Peterman et al. 2011, pp. 1486).

The first strand of the literature encompasses the overwhelming majority of the empirical studies on the topic. These studies generally use the gender of the head of household as the main explanatory variable to identify the gender gap. The common assumptions of these research efforts are that the members of a given household do not necessarily differ in their sex, age, productive capacity and/or personality profiles; that information is shared symmetrically between cooperative individuals; and that differences in the quantity and quality of land and non-land inputs used by different individuals within or across study households are negligible (Schultz, 2001; Peterman et al., 2011). The extent to which these assumptions are valid in a given sub-Saharan African setting depends on (i) the complexity of familial structures, including monogamous, polygamous, skipped-generation, and multi-generation households (Peterman et al., 2011), and (ii) the persistence of rights and obligations that affect men and women differently and that are underscored by biological differences, social and religious norms, and customs that jointly dictate the division of labor, land, and proceeds from production units (Saito et al., 1994).

A considerable majority of the studies of the second strand of the literature on the gender gap in sub-Saharan Africa originate from West Africa, specifically from Ghana and Burkina Faso, where it is common for households to have several agricultural plots and for male and female plot managers to coexist in study households. This allows authors to control for unobserved time-invariant household-crop-level heterogeneity in a multivariate regression framework and to estimate agricultural production functions for plots cultivated with the same crop, managed or owned by men and women in the same household. As such, they have evolved to be the most influential studies on gender differences in agricultural productivity in sub-Saharan Africa, documenting the potential Pareto-inefficient nature of within-household allocation of productive resources. The gender gap is typically identified by the magnitude and statistical significance of the regression coefficient associated with the gender of the plot manager/owner.

The evidence from the second strand indicates that in some contexts, descriptive mean differences in agricultural productivity across plots owned/managed by males vs. females continue to be large and statistically significant in multivariate analyses that control for differences in input use (Saito et al., 1994 for Nigeria; Udry, 1996 for Burkina Faso; Quisumbing et al., 2001 for Ghana; Peterman et al., 2011 for Uganda), while in other contexts, the gender gap ceases to be statistically significant once the researcher controls for differential utilization of productive inputs (Saito et al., 1994 for Kenya;

Gilbert et al., 2002 for Malawi; Akresh, 2005 for Burkina Faso; Goldstein and Udry, 2008 for Ghana).[9]

A common limitation of these studies is the reliance on data that are at best regionally-representative in terms of population dynamics (with the exception of Akresh, 2005), whereby results have limited external validity beyond the study area, within or across countries. For instance, Akresh (2005) uses nationally-representative data from Burkina Faso, and while he is able to replicate the findings of Udry (1996) by focusing on a subset of villages that are in close proximity to the areas underlying Udry's analysis, Akresh cannot recover the same relationships based on the data collected in other parts of the country. This discrepancy highlights the importance of revisiting the body of evidence on the gender gap in sub-Saharan Africa by using nationally-representative data.

Another limitation observed in the second strand of the relevant literature is the disproportionate focus on West Africa. It is important to investigate the extent and correlates of the gender gap in alternative sub-Saharan African settings with different sets of rights and obligations that differently affect the distribution of productive resources across men and women. Finally, in the case of empirical studies that document statistically insignificant differences in agricultural productivity between female-managed and male-managed plots, conditional on plot-level observable and household-level unobservable attributes, the analytical framework is not set up to isolate relative

[9] See Peterman et al. (2011) for a succinct review of the main findings of the studies cited here. Only Udry (1996), Quisumbing et al. (2001), Akresh (2005), Goldstein and Udry (2008), and Peterman et al. (2011) conduct within-household analysis.

contributions of relevant attributes towards the observed gender gap for the purpose of prioritizing areas for policy interventions.

3 Malawi: Agricutural Productivity and Gender

3.1 The Country Context[10]

Malawi is a small, population-dense, land-locked country in Southern Africa, with 94,080 square kilometers of land. The 2010 mid-year population projection and annual population growth rate stand at 14.5 million persons and 3.25 percent, respectively, and 85 percent of the population reside in rural areas (NSO, 2012). Agriculture is not only the backbone of Malawi's economy but also an essential part of its social fabric. The sector accounts for 30 percent of the Gross Domestic Product (GDP), and 84 percent of Malawian households own and/or cultivate land[11]. The production system is overwhelmingly rainfed, characterized by limited access to irrigation and diminishing average land holding sizes due to population pressures. The rainfall is unimodal, and maize is the main staple crop, grown by nearly 100 percent of the farming household population.[12]

[10] Unless otherwise stated, the statistics reported in Section 3.1 originate from data.worldbank.org/country/Malawi.

[11] The GDP contribution of agriculture is for 2011. The estimate of the percentage of Malawian households owning and/or cultivating land is based on the Third Integrated Household Survey (IHS3) data.

[12] The estimate is based on the IHS3 data.

Over the last two decades, agricultural productivity, as measured by maize yields (kilogram/hectare), has been erratic, as shown in Figure 1. The factors that are commonly cited as underlying the agricultural productivity trend include weather variability, declining soil fertility, limited use of improved agricultural technologies and sustainable land management practices, rationed agricultural extension services, market failures, and underdeveloped and poorly maintained infrastructure (World Bank, 2007). The majority of the farming households still practice subsistence agriculture: the rates of market participation among farming households in general and maize-producing households in particular are 42 and 15 percent, respectively.[13] The inconsistent agricultural performance has direct implications for living standards, given the predominantly rural nature of the country and its heavy reliance on agriculture.

Poverty remains widespread and persistent, particularly among female headed households. Based on the data from the Second Integrated Household Survey (IHS2) 2004/05 and the Third Integrated Household Survey (IHS3) 2010/11, the national absolute poverty rate of 52.4 percent in 2004/05 declined only marginally to 50.7 percent in 2010/11. The trends in rural poverty followed a similar pattern: a rate of 55.9 percent in 2004/05 vs. 56.6 percent in 2010/11.[14] Focusing on the gender dimensions of poverty, while the absolute poverty rate among male-headed households was estimated at 49 percent in 2010/11, the comparable figure among female-headed households was 57 percent. In an effort to combat poverty and boost national food security, the Malawian

[13] The estimate is based on the IHS3 data.

[14] The difference between the IHS2 and the IHS3 national absolute poverty rates is not statistically significant. The IHS3 rural poverty rate is also statistically indistinguishable from its IHS2 counterpart.

Government has embarked on an ambitious annual fertilizer and seed subsidy program known as the Farm Input Subsidy Program (FISP), starting with the 2005/06 agricultural season. During the 2009/10 agricultural season (the reference agricultural season for over 75 percent of our sample), close to 50 percent of the farming household population is estimated to have participated in the program.[15] Stagnant poverty levels raise questions on the effectiveness of the FISP in alleviating poverty and food insecurity in a sustainable fashion, which should be subject to further empirical investigation.[16]

3.2 Data

This study uses data from the Third Integrated Household Survey (IHS3), collected from March 2010 to March 2011 by the Malawi National Statistical Office, with support from the World Bank Living Standards Measurement Study - Integrated Surveys on Agriculture (LSMS-ISA) project.[17] The IHS3 data were collected within a

---

[15] The exchange rate for the IHS3 period is MK150 = US$1. During the 2009/10 agricultural season, each FISP beneficiary was entitled to vouchers that allowed them to purchase (i) two 50 kilogram bags of maize fertilizer at 500 Malawi Kwacha (MK) per bag, (ii) either 3 kilograms of hybrid maize seed or 10 kilograms of open-pollinated variety maize seed for the commercial market value net of the 1500 MK subsidy from the Government, (iii) 200 grams of storage pesticide for 100 MK, and (iv) 1 kilogram of legume seed (groundnuts, soybeans, beans and pigeon peas) for free. Upon the allocation of vouchers across the districts and the villages within each district, the program relies on community-based targeting to identify beneficiaries at the local-level, and is supposed to target households that are (i) resource poor, (ii) permanent village residents, and (iii) own and cultivate land, with preference given to heads of households that may be female, orphan, elderly, physically-challenged, or HIV-positive or individuals that look after the elderly and physically-challenged (MoAFS, 2009).

[16] Concerns regarding the effectiveness of FISP in reducing poverty and achieving sustainable gains in maize production have been raised (Ricker-Gilbert and Jayne, 2011; Holden and Lunduka, 2010). More recently, Ricker-Gilbert and Jayne (2012) focus on the question of whether FISP can simultaneously boost maize production and reduce poverty, and document that major returns from subsidized fertilizer accrue almost exclusively to households at the top of the maize production and value of total crop output distributions.

[17] The IHS3 data and documentation are publicly available through the LSMS website (www.worldbank.org/lsms).

two-stage cluster sampling design, and are representative at the national, urban/rural, regional, and district levels, covering 12,271 households in 768 enumeration areas (EAs). The IHS3 instruments included Household, Agriculture, Fishery, and Community Questionnaires.

All sample households were administered the multi-topic Household Questionnaire that collected individual-disaggregated information on demographics, education, health, wage employment, nonfarm enterprises, anthropometrics, and control of income from non-farm income sources, as well as data on housing, food consumption, food and non-food expenditures, food security, and durable and agricultural asset ownership, among other topics. The sample households that were involved in agricultural activities (through ownership and/or cultivation of land, and/or ownership of livestock) were administered the Agriculture Questionnaire. The Agriculture Questionnaire solicited information on land areas, physical characteristics, labor and non-labor input use, and crop cultivation and production at the plot level, separately for the reference rainy and dry seasons. [18] The data allow for agricultural production estimates at the plot level and for the identification of the manager of the plot[19], as well as household members that owned[20] and/or worked on each plot.[21] Handheld global positioning system (GPS)-based

[18] A plot was defined as a continuous piece of land on which a unique crop or a mixture of crops is grown, under a uniform, consistent crop management system, not split by a path of more than one meter in width. Plot boundaries were defined in accordance with the crops grown and the operator.

[19] For each plot, the following question was asked to identify the primary decision maker/manager: "Who in this household makes the decisions concerning crops to be planted, input use and the timing of cropping activities on this plot?" The questionnaire allowed for identification of one manager per plot, on whom individual-level information could be recovered from the Household Questionnaire.

[20] For each plot, the following question was asked to identify the plot owners: "Who owns this plot?" The question allowed up to 2 household members to be specified as owners.

locations and land areas of the plots were recorded, permitting us to link household- and plot-level data to outside geographic information system (GIS) databases.

The descriptive statistics and the results from the tests of mean differences by the gender of the plot manager are presented in Table 1. The full sample consists of 16,372 plots, 26% of them managed by females.[22] Table 1 clearly demonstrates the (unadjusted) gender gap: the average gross value of output per hectare, our proxy for agricultural productivity, is 25% lower for the female-managed plot sample.[23] The gender differences in agricultural productivity are also evident in the comparison of the Kernel density estimates of the log of gross value of output per hectare for male- and female-managed plots, as displayed in Figure 2. The overwhelming majority of the differences in the average values of the observable covariates across male- vs. female-managed plots in Table 1 are statistically significant at the 1 percent level. For the purposes of the ensuing

[21] 81 percent of the plots in our sample are reported to be owned. Among the owned plots, 15 percent have joint ownership, of which the predominant form is male-female. The remaining 38 percent and 47 percent of the owned plots are under sole male ownership and sole female ownership, respectively.

[22] The IHS3 identified 18,917 plots that were reported to have been owned and/or cultivated during the reference rainy season (2008/09 or 2009/10). 618 plots are not considered for analysis since they lacked either GPS-based plot coordinates or GPS-based plot area. 1,314 plots are dropped since they are either fallow or missing production information. 199 plots are not included in the sample since unit values could not be computed reliably for at least one of the crops reported to be cultivated on the plot. 11 plots do not have a manager identified and 67 plots have at least one missing value among the independent variables of interest. Finally, top and bottom 1 percent of the distribution of the log of gross value of output per hectare are trimmed, corresponding to 336 plots. These exclusions leave us with the final analysis sample of 16,372 rainy season plots.

[23] The plot-level gross value of output in Malawi Kwacha (MK) is calculated by first multiplying the kilogram-equivalent quantity of production for each crop on a given plot by the median crop sales value per kilogram within the corresponding EA, and then aggregating across values of crop production. The median crop sales value per kilogram is computed within the corresponding EA only if at least 10 values are available from the survey data. Otherwise, the median crop sales value per kilogram is computed at a higher level, in the order of traditional authority, district, region, and country. Our outcome variable is computed by normalizing plot-level gross value of output with GPS-based cultivated plot area.

discussion, we focus on the differences that are statistically significant at least at the 5 percent level.

The incidence of manager-head of household correspondence is 99 percent for the male-managed plot sample, while the analogous statistic is 80 percent for the female-managed plot sample. Female-managed plots are, on average, overseen by individuals that are 5 years older and have 2 less years of schooling with respect to their male-managed comparators. A significantly higher percentage of female-managed plots exhibit manager-owner correspondence (77 vs. 58 percent) and 75 percent of the female-managed plot sample are exclusively female-owned, featuring either a sole female owner or dual female owners within the household.[24] The incidences of joint ownership and exclusive-male ownership stand at 4 and 3 percent, respectively, among female-managed plots. In comparison, male-managed plots are distributed more evenly across the ownership categories of exclusive-male (43 percent), exclusive-female (23 percent), and joint male-female (15 percent).

Although the average GPS-based plot area is 0.39 hectare, female-managed plots are, on average, 12 percent smaller than their male-managed counterparts. The use of inorganic fertilizer is lower on female-managed plots, whether measured by incidence, average unconditional amount per hectare, or average conditional amount per hectare. These trends may signal gender differences in FISP fertilizer voucher distribution and

[24] The overwhelming majority of the owned plots (83 percent) are acquired through inheritance. Another 12 percent is reported to have been granted by local leaders. The remaining are acquired as bride price (2 percent), purchased with title (1 percent) and purchased without title (1 percent).

redemption outcomes. In fact, the IHS3 data indicates that even though female-headed households are just as likely to receive a fertilizer voucher as their male-headed counterparts, conditional on receipt, the *average number* of fertilizer vouchers that are received among female-headed households (1.56) is lower than the analogous statistic for male-headed households (1.63), and the difference is statistically significant at the 1 percent level. Similarly, conditional on receipt, the average number of fertilizer vouchers that are *redeemed* stands at 1.57 and 1.48 for male-headed and female-headed households, respectively, and the difference is again statistically significant at the 1 percent level.[25]

In terms of household labor use, the dynamics are drastically different on female-managed plots *vis-à-vis* their male-managed comparators, as can be seen in Table 1. Although the average incidence, average unconditional amount, and average conditional amount of household adult male labor input per hectare are significantly higher on male-managed plots, the opposite is true concerning household adult female, household child, and exchange labor use on female-managed plots.[26] The relatively higher household child labor and exchange labor input on female-managed plots might be possible responses to being rationed out of household adult male labor. Furthermore, Table 1 shows statistically different cultivation patterns by gender of the plot manager, with female-

[25] The factors behind this pattern are being investigated further as part of a parallel research program on the FISP beneficiary targeting performance and productivity impacts.

[26] Adult is defined as being at least 15 years of age. The plot-level measures of household adult male, adult female, and child labor input are the summations of rainy season labor hours across household members reported to have worked on a given plot. Individual labor input is computed as the multiplication of the number of weeks a household member worked on a given plot during the reference rainy season, the typical number of days worked per week during the reported number of weeks, and the typical number of hours worked per day during the reported number of weeks. The plot-level measure of hired labor (exchange) input is the sum of aggregate men, women, and child hired (exchange) labor days.

managed plots exhibiting a higher incidence of intercropping and male-managed plots recording, on average, higher shares of plot area (i) under improved seeds (mainly maize, complemented by groundnuts and rice) and (ii) under export crops (mainly tobacco, complemented by cotton). Female-managed plots are also 4 percentage points less likely to be associated with households that receive agricultural extension service on topics that relate to crop production and marketing. Lastly, male-managed plots are, on average, more likely to be associated with households with higher levels of wealth and access to agricultural implements.[27]

Table 2 presents the naïve plot-level regression results on the gender gap, where the dependent variable is the log of gross value of output per hectare. The findings presented in columns 1, 2 and 3 originate from regressions that, in addition to the dummy variable on female plot management, control only for agro-ecological zone, regional, and district fixed-effects, respectively. The gender gap estimates range from 22 to 25 percent. These results indicate a statistically and economically large difference between male and female farmers. In what follows, we seek to understand the factors associated with this gap.

[27] The household wealth index is constructed using principal component analysis, and takes into account the number of rooms in the dwelling, a set of dummy variables accounting for the ownership of (i) dwelling, (ii) mortar, (ii) bed, (iii) table, (iv) chair, (v) fan, (vi) radio, (vii) tape/CD player, (viii) TV/VCR, (ix) sewing machine, (x) paraffin/ kerosene/ electric/ gas stove, (xi) refrigerator, (xii) bicycle, (xiii) car/motorcycle/minibus/lorry, (xiv) beer brewing drum, (xv) sofa, (xvi) coffee table, (xvii) cupboard, (xviii) lantern, (xix) clock, (xx) iron, (xxi) computer, (xxii) fixed phone line, (xxiii) cell phone, (xxiv) satellite dish, (xxv) air-conditioner, (xxvi) washing machine, (xxvii) generator, (xxviii) solar panel, (xxix) desk, and a vector of dummy variables capturing access to improved (i) outer walls, (ii) roof, (iii) floor, (iv) toilet, and (v) water source. The household agricultural implement access index is also computed using principal components analysis, and covers a range of dummy variables on the ownership of (i) hand hoe, (ii) slasher, (iii) axe, (iv) sprayer, (v) panga knife, (vi) sickle, (vii) treadle pump, (viii) watering can, (ix) ox cart, (x) ox plough, (xi) tractor, (xii) tractor plough, (xiii) ridger, (xiv) cultivator, (xv) generator, (xvi) motorized pump, (xvii) grain mail, (xviii) chicken house, (xix) livestock kraal, (xx) poultry kraal, (xxi) storage house, (xxii) granary, (xxiii) barn, and (xxiv) pig sty.

Table 3 provides an additional estimate of the gender gap, but now conditional on additional covariates commonly found in the literature (Peterman et al., 2011). Column 1 presents the results from a pooled regression that includes both male- and female-managed plots. Once we control for key factors of production, the gender gap is reduced to 4.5 percent and is now statistically significant only at the 10 percent level. In the end, this type of analysis does not allow us to delve deeper into the process that underlies the movement from the unconditional gender gap of 25.4 percent to the conditional gender gap of 4.5 percent. In the following section, we apply a decomposition approach that will allow us to unpack the relative contributions of different factors towards this gap and to suggest priority areas for policy interventions.

4 Mean Decomposition Methodology

Regression-based decomposition methods have been widely utilized in labor economics following the seminal papers of Oaxaca (1973) and Blinder (1973), notably as part of the analyses of the gender wage gap, union wage gap, and growing wage inequality (O'Neill & O'Neill, 2006, Fortin, 2006). Despite the extensive use of Oaxaca-Blinder regression-based mean decomposition among applied economists over the last three decades and the advances that have been made to extend the application to the decomposition of distributional statistics besides the mean, the questions attempted to be addressed by the method require a strong set of assumptions (Fortin et. al., 2011).

In particular, these methods follow a partial equilibrium approach, where observed outcomes for one group can be used to construct various counterfactual scenarios for the other group (Fortin et. al., 2011). Another limitation is that while decompositions are useful for quantifying, purely in an accounting sense, the contribution of various factors to a difference in an outcome across groups or a change in an outcome for a particular group over time, they are based on correlations, and hence cannot be interpreted as estimates of underlying causal parameters (Fortin et. al., 2011). However, decomposition methods do document the relative quantitative importance of factors in explaining an observed gap, thus suggesting priorities for further analysis and, ultimately, policy interventions (Fortin et. al., 2011).

To document the extent and drivers of the gender gap in Malawi, we first rely on an Oaxaca-Blinder regression-based mean decomposition. We assume the log of an agricultural productivity measure (Y), namely gross value of agricultural output per hectare, for male- (M) and female- (F) managed plots estimated as:

(1) $Y_G = \beta_{G0} + \sum_{k=1}^{K} X_{Gk}'\beta_{Gk} + \varepsilon_G$

where *G* indicates the gender of the plot manager; *X* is a vector of k observable, plot-, household- and/or community-level explanatory variables; β is the associated vector of intercept and slope coefficients; and ε is the error term under the assumption that E(ε_M) = E(ε_F) = 0.

The *gender gap* "*D*" is expressed as the mean outcome difference:

(2) $D = E(Y_M) - E(Y_F)$.

Equations (1) and (2) imply that:

(3) $E(Y_M) = E(\beta_{M0} + \sum_{k=1}^{K} X_{Mk}\beta_{Mk} + \varepsilon_M) = \beta_{M0} + \sum_{k=1}^{K} E(X_{Mk})\beta_{Mk}$

(4) $E(Y_F) = E(\beta_{F0} + \sum_{k=1}^{K} X_{Fk}\beta_{Fk} + \varepsilon_F) = \beta_{F0} + \sum_{k=1}^{K} E(X_{Fk})\beta_{Fk}$

and, Equation (2) could be rewritten as:

(5) $D = E(Y_M) - E(Y_F) = \beta_{M0} + \sum_{k=1}^{K} E(X_{Mk})\beta_{Mk} - \beta_{F0} - \sum_{k=1}^{K} E(X_{Fk})\beta_{Fk}$.

Subsequently, we define β* as the vector of coefficients that is obtained from a regression of Y that is based on the pooled plot sample and includes the group membership identifier, i.e. a dummy variable identifying female-managed plots. The inclusion of the group membership indicator in the pooled regression for the estimation of β* takes into account the possibility that the mean difference in plot-level productivity measure is explained by gender of the plot manager, avoiding a possible distortion of the decomposition results due to the residual group difference reflected in β* (Jann, 2008). Rearranging Equation (5) by adding and subtracting (i) the slope coefficient of the pooled regression (β_0^*), and (ii) the return to the observable covariates of each group valued at β^* ($X_{Mk}\beta_k^*$ and $X_{Fk}\beta_k^*$), we obtain:

(6)

$$D = \underbrace{\sum_{k=1}^{K}[E(X_{Mk}) - E(X_{Fk})]\beta_k^*}_{Component\ 1:\ Endowment\ Effect} \quad +$$

$$\underbrace{\underbrace{(\beta_{0M} - \beta_0^*) + \sum_{k=1}^{K}[E(X_{Mk})(\beta_{Mk} - \beta_k^*]}_{Male\ Structural\ Advantage} + \underbrace{(\beta_0^* - \beta_{0F}) + \sum_{k=1}^{K}[E(X_{Fk})(\beta_{Fk} - \beta_k^*]}_{Female\ Structural\ Disadvantage}}_{Component\ 2:\ Structure\ Effect}$$

where $\beta_{M0}, \beta_{F0}, \beta_0^*, \beta_{Mk}, \beta_{Fk}, \beta_k^*$ (k=1….K) are the estimated intercept and slope coefficients of each covariate included in the regressions for the male-managed, female-managed and pooled plot samples.

Equation (6) is known as the *aggregate decomposition*. The first component is the *endowment effect*, i.e. the portion of the gender gap that is explained by differences in the levels of observable covariates between both groups. It is simply the sum across all covariates, of the differences by group, valued at the corresponding "average" return. The second component is the *structure effect,* i.e. the portion of the gender gap driven by deviations of each group's return from the corresponding "average" return. The first term of the structure effect $(\beta_{0M} - \beta_0^*) + \sum_{k=1}^{K}[E(X_{Mk})(\beta_{Mk} - \beta_k^*]$ represents the *male* *structural advantage*, which is equal to the portion of the gender gap accounted for by deviations of male regression coefficients from pooled counterparts. The second term of the structure effect $(\beta_0^* - \beta_{0F}) + \sum_{k=1}^{K}[E(X_{Fk})(\beta_{Fk} - \beta_k^*]$ represents the *female structural disadvantage*, which is equal to the portion of the gender gap driven by deviations of pooled regression coefficients from female counterparts.[28]

In practice, we estimate equation 1 for (i) male-managed plots, (ii) female-managed plots, and (iii) the pooled plot sample (with a dummy variable identifying female-managed plots), and use the resulting vector of coefficients β_M, β_F, and β^*, together with the mean values for each covariate for each group X_M and X_F to compute the components of equation (6). Moving beyond the aggregate decomposition, the detailed decomposition involves subdividing the endowment and structure effects into the respective contributions of each observable covariate, which correspond to the variable-specific subcomponents of the summations included in equation (6).

Fortin et al. (2011) present a detailed account of the assumptions required to identify the population parameters of interest. Two crucial assumptions for the validity of aggregate decomposition are (i) overlapping support and (ii) ignorability. *Overlapping support* implies that no single value of $X = x$ or $\varepsilon = e$ exists to identify female plot

[28] The use of the term "disadvantage" is tied to the subsequent section's discussion of the regression coefficients estimated from the pooled, male-managed, and female-managed plot samples. With respect to their counterparts estimated from the pooled plot sample, the regression coefficients from the female-managed plot sample that are expected to be positive and that are associated with key factors of production are consistently positive but lower in absolute terms. Conversely, the use of the term "advantage" is linked to the same set of regression coefficients being higher in the male-managed plot sample with respect to those from the pooled plot sample.

management. *Ignorability* refers to the random assignment of female plot management conditional on observable attributes. The additional essential assumptions required by detailed decomposition to identify the individual contribution of each covariate include *additive linearity* and *zero conditional mean*. The latter implies that ε is independent of X. In other words, we assume that there is no unobservable heterogeneity that jointly determines the outcome and observable attributes. It should be noted that even if the additional assumptions required by detailed decomposition may not hold true, aggregate decomposition would remain valid as long as overlapping support and ignorability assumptions are tenable.

In exploring the existence and extent of the gender gap in a multivariate framework, the validity of findings largely depend on the plausibility of ignorability and zero conditional mean assumptions, i.e. the extent to which the identification strategy addresses possible unobservable household-/plot-level heterogeneity that jointly determines plot agricultural productivity and observable covariates, including whether a plot is managed by a female. While the most rigorous studies on the gender gap recognize the need for an instrumental variable strategy to deal with potentially endogenous observables, recovering instrumental variables that predict endogenous covariates without directly influencing the outcome is often not possible. A subset of the studies that are reviewed in Section 2 and that feature plot-level analyses have attempted to address the potential bias by controlling for direct measures of plot soil quality and household-crop fixed effects. With panel data, time fixed effects have also been included in the specifications.

Furthermore, to deal with the possibility that the male and female plots might be physically systematically different from each other along the dimensions that are finer than the observable variations in soil physical infrastructure and quality, an alternative identification strategy has been to rely on a spatial fixed effects estimator that allows for local neighborhood effect in unobserved land quality that could be correlated with the gender of plot manager and the other regressors. The spatial fixed effects are differenced out by modeling the difference between the plot-level outcome and the average of the outcome across plots from other households within a critical distance as a function of a vector of plot-/household-level variables that are differenced from their matched plot-/household-level neighborhood averages (see Goldstein and Udry (2008) for an example of this).

We lack, in our case, plot-level measures of soil quality, and the nature of farm organization as captured in the IHS3 data does not allow us to feature household fixed effects as a central piece of our empirical strategy. The average number of plots cultivated by Malawian farming households is 1.76, significantly less than the comparable statistics from West African settings that have largely informed the analysis of the gender gap in sub-Saharan Africa thus far. The managers identified across agricultural plots cultivated by a given household also correspond to the head of household in an overwhelming sample of households that report to be cultivating multiple plots. Specifically, there are only 109 households that cultivate multiple plots and exhibit within-household variation in terms of the gender of the plot managers, corresponding to 1.7 percent of our plot sample. Nevertheless, we attempt to lend as much support to the overlapping support, ignorability, and zero conditional mean assumptions as possible by relying on all available data and econometric methods at our disposal. These sensitivity analyses are presented later in Section 5.3.

5 Mean Decomposition Results

The first step in the mean decomposition is the estimation of equation (1). This is done separately for the pooled, male-managed and female-managed plot samples, and the results reported in Table 3, Columns 1, 2 and 3, respectively.

We find that the log of GPS-based plot area has a negative coefficient that is statistically significant at the 1 percent level in each plot sample. This finding is consistent with recent studies that have investigated and provided support for the inverse yield hypothesis (see Larson et al., 2012 and the references cited therein). A key variable that is positively associated with the log of gross value of output per hectare, irrespective of the plot sample, is the log of inorganic fertilizer use per hectare. However, the return to inorganic fertilizer use (i.e. the coefficient) is higher within the male-managed plot sample in comparison to the female-managed plot sample, and this difference is statistically significant.

The log of household adult male labor hours per hectare has a sizeable and positive coefficient that is statistically significant at the 1 percent level within the male-managed plot sample, while the comparable estimate within the female-managed plot sample is not statistically significant. In contrast, the log of household adult female labor hours per hectare has a positive and statistically significant coefficient across both plot samples, albeit a larger effect, in terms of both magnitude and statistical significance, among female-managed plots.

The coefficients for the shares of plot area under improved seeds and under export crops have sizably positive and statistically significant coefficients at the 1 percent level across all plot samples of interest. Conversely, the child dependency ratio, which is defined as the number of household members below the age of 10 divided by the number of household members aged 10 years and above, has a substantial negative coefficient that is statistically significant at the 1 percent level only within the female-managed plot sample. The comparable statistics for the pooled and male-managed plot samples are not statistically significant.

In addition, although household size has a positive coefficient that is statistically significant irrespective of the plot sample, the magnitude of the coefficient within the female-managed plot sample is three times larger than within the male-managed plot sample. The gender differences in returns to household size and child dependency ratio imply that the burden of childcare is more likely to reduce female agricultural productivity.

The decomposition of the mean gender gap, which is estimated at 25.4 percent, is presented in Table 4. Panel B presents the aggregate decomposition components, namely the endowment effect, the male structural advantage, and the female structural disadvantage. Panel C includes the results from the detailed decomposition, whereby a positive coefficient suggests that the relevant covariate contributes positively to increasing the gender gap.

5.1 Aggregate Decomposition

The aggregate decomposition indicates that the endowment effect (20.9 percentage points), i.e. the portion of the gender gap driven by gender differences in levels of observable attributes, accounts for 82 percent of the mean gender differential in agricultural productivity. The female structural disadvantage is estimated at 4.5 percentage points, explaining the remaining 18 percent of the gender gap. The aggregate decomposition reinforces the notion that large and significant gender disparities in access to inputs and in asset ownership are central factors behind the gender gap.

5.2 Detailed Decomposition

The detailed decomposition of the endowment effect is reported in Table 4, Panel C, Column 1. As noted above, the estimates are a function of the mean differences reported in Table 1 by the gender of the plot manager, and the pooled regression coefficients reported in Table 3. The percentage contributions that are noted below should be understood as correlations, rather than causal parameters, and are obtained by dividing the coefficient in question either by the endowment effect (0.209) or by the gender gap (0.254).

In Section 3.2, we noted that male-managed plots tend to be overseen by individuals that have higher years of schooling and who originate from larger and wealthier households that access agricultural extension more frequently. Male-managed plots also exhibit higher (i) incidence of pesticide use, (ii) inorganic fertilizer use per hectare, (iii) household adult male labor input per hectare, (iv) share of plot area under improved seeds, and (v) share of plot area under export crops. In view of the positive correlation with these covariates and agricultural productivity, we find these variables to be contributing positively towards the endowment effect, thereby widening the gender gap. Conversely, the higher rate of household adult female labor and exchange labor provision within the female-managed plot sample, as well as the positive association between these covariates and agricultural productivity imply that these variables contribute negatively towards the endowment effect, hence working to close the gender gap. The smaller plot areas farmed by female managers also appear to be a contributing factor in shrinking the gender gap given that in these data, there is an inverse relationship between cultivated plot area and agricultural productivity.

The factors that comprise the majority of the endowment effect are the log of household adult male labor hours per hectare and the share of plot area under export crops. The covariates explain 46 percent and 40 percent of the endowment effect, and account for 38 percent and 33 percent of the gender gap, respectively. The positive contributions of the other covariates towards the endowment effect (and the gender gap in parenthesis) are as follows: (i) 14 percent (11 percent) for the household agricultural implement access index, (ii) 9 percent (7 percent) for the household wealth index, (iii) 8 percent (6 percent) for manager years of schooling, (iv) 8 percent (6 percent) for the log of plot inorganic fertilizer use per hectare, (v) 6 percent (5 percent) for household size, and (vi) 2 percent (2 percent) for share of plot area under improved seeds. The negative contributions of the aforementioned covariates towards the endowment effect (and the gender gap in parenthesis) are as follows: (i) 24 percent (20 percent) for the log of GPS-based plot area and its squared term combined, (ii) 8 percent for logged household female adult labor hours per hectare, and (iii) 2 percent (2 percent) for the log of exchange labor days per hectare.

The detailed decompositions of the male structural advantage and the female structural disadvantage are presented in Columns 2 and 3 of Table 4, Panel C. The coefficients that are large and statistically significant signal differential treatments of male vs. female plot manager by markets, formal institutions, and informal social institutions. Findings related to inorganic fertilizer use, plot measures of household adult male and adult female labor provision, household size, and child dependency ratio are noteworthy.

First, it is not only the differences in the inorganic fertilizer endowment that contribute to the gender gap, but also the relatively higher return to inorganic fertilizer among the male-managed plots in comparison to their female-managed counterparts. The same applies to the log of household adult male labor hours per hectare. The underlying causes of this finding are the subject of future research but may indicate household adult male labor supervision difficulties on female-managed plots.

The fact that household adult male labor input is associated with a wider gender gap is, however, partially offset by the higher returns that household adult female labor provides on female-managed plots. Regarding the child dependency ratio, although the contribution of this factor towards the endowment effect is zero, its contribution towards the female structural disadvantage is large and positive, driven by the sizeable and highly significant negative association between this variable and agricultural productivity solely within the female-managed plot sample. This result highlights the differential productivity impacts of heterogeneous household roles assumed by male and female managers. Since female managers, who are just as likely to be household heads or spouses, are more likely to combine farm management with household duties, including child care, their pattern of time use is directly related to their low productivity outcomes.

5.3 Sensitivity Analyses

As noted earlier, the crucial assumptions for the validity of the aggregate decomposition include overlapping support and ignorability. The key assumptions additionally required by the detailed decomposition are additive linearity and zero conditional mean. A methodology that is proposed by Imbens and Rubin (2009) to assess the feasibility of the overlapping support assumption is centered on the idea of calculating a scale-free normalized difference for each covariate. They assert that the overlapping support across the groups of interest, in our case female- vs. male-managed plots, is adequate if the scale-free normalized differences across the covariates are less than 0.25. Table H.1 in Appendix H presents the scale-free normalized difference of the variables used in the regressions. Only 2 out of 29 independent variables have a normalized difference greater than 0.25.

In trying to lend support to ignorability and zero conditional mean assumptions, we use all available data and econometric tools at our disposal, and first rely on an empirical approach that was pioneered by Altonji (1988), Murphy and Topel (1990), and Altonji et al. (2005), based on the idea that the amount of selection on observable variables provides a guide to the extent of selection on unobservable counterparts. We use an informal version of the methodology applied by Acemoglu et al. (2001) and Altonji et al. (2005), and incorporate into our base specification, in a phased-in fashion, thematically-grouped control variables such that each regression is estimated with a different set of additional independent variables and that the results are compared to those from the base specification. Our purpose is to gauge the stability of the key regression coefficients that underlie our decomposition results. If the coefficients on the covariates included in the base specification, including the female plot management dummy in the pooled regression, are stable subsequent to the incorporation of additional covariates, they are less likely to change if we are able to take into account potentially missing omitted variables.

To perform this analysis, we consider the following sets of variables: (i) district fixed effects, (ii) plot geospatial characteristics informed by GIS data, (iii) other plot characteristics solicited by the IHS3, (iv) additional household characteristics, and (v) additional community characteristics. Table H.2 in Appendix H includes the detailed list of the variables included in each set. Tables H.3, H.4 and H.5 present the base regression results and the estimates from the regressions including the additional controls for the pooled, male-managed, and female-managed plot samples, respectively. An overwhelming majority of the coefficients, with respect to the base specification, are stable across the specifications and the plot samples, and do not change sign or significance. This suggests that the assumptions of ignorability and zero conditional mean might not be unfounded.

As part of the sensitivity analyses, we also conduct our analysis on the subset of plot observations from households in which male and female plot managers coexist. The size of this sample is 292 plots (approximately 2 percent of our sample) originating from 109 households. The concern motivating this analysis is that there might be unobserved household characteristics that might jointly determine productivity outcomes and within-household assignment of plots to managers, which may be biasing our estimates. Table H.6 in Appendix H presents the results from pooled regressions that use the aforementioned sample of 292 plots, and that are compared with the estimates from the base regression that is informed by the entire pooled sample (Column 1). Column 2 includes the coefficient estimates from a regression that is identical to the base specification but is estimated using the reduced plot sample, and Column 3 presents the findings from a regression that is fit among the same sample of plots but with household fixed effects incorporated in the base specification.

Although the coefficients associated with female plot management in Columns 2 and 3 are not statistically significant (likely due to the small sample), they are economically relevant, with values of 6.6 percent and 11.4 percent, respectively. Furthermore, Table H.7 in Appendix H presents the decomposition of the mean gender gap using the sample of 292 plots, but informed by the regression set-up that is identical to the base specification. The mean gender gap is equal to 27.4 percent, and the differential is mostly explained by the endowment effects of the inorganic fertilizer use and the share of plot area under export crops. These results are consistent with the findings presented earlier.

Finally, we replicate the entire analysis by using the plots cultivated with maize and the log of maize production per hectare as an alternative proxy for agricultural productivity. These results are available upon request and are strongly in line with the findings reported thus far. As shown in Column 1 of Table 5, the gender gap in maize yields is equal to 22.4 percent, and approximately three-quarters of the observed differential are driven by the endowment effect. The similarity of the results supports the hypothesis that the independent variables in our base specification that uses the log of plot-level gross value of output per hectare as the dependent variable capture the possible bias that unobserved technology choice parameters may otherwise cause.

6 Recentered Influence Function (RIF) Decomposition

Our decomposition findings suggest that more than 80 percent of the mean gender gap is explained by differences in observable covariates, and the direct pay-off to addressing market and institutional failures that affect men and women differentially is economically significant. While it is important to show this with nationally-representative data, going beyond the "average" farmer and understanding the *heterogeneity* in constraints faced by farmers with different gender and productivity profiles is crucial for the design and implementation of better targeted interventions aimed at bridging the gap. An important question is whether our findings, which are based on the sample means, are robust to the decomposition of alternative distributional statistics beyond the mean.

A method that is similar in spirit to the mean decomposition uses the recentered influence function (RIF) regressions proposed by Firpo et al. (2009) and provides a straightforward framework within which across-group differences in any distributional statistic could be decomposed. We rely on the RIF decomposition to provide estimates of the aggregate and detailed decomposition of the gender gap at different percentiles of the agricultural productivity distribution.

A RIF regression is similar to a standard OLS regression, except that the dependent variable, Y, is replaced by the RIF of the distributional statistic of interest. The approach assumes that the conditional expectation of the $RIF(Y;v)$ can be modeled as a linear function of observable attributes, X, such that $E[RIF(Y;v)|X] = X\gamma$, as in the mean decomposition. Assuming that $IF(y; v)$ is the influence function corresponding to an observed productivity outcome y, for the distributional statistic $v(F_Y)$, the RIF is defined as:

(7) $RIF(y; v) = v(F_Y) + IF(y; v).$

In the case of quantiles, the influence function is equal to:

(8) $IF(Y; Q_T) = \frac{(T - \mathbf{1}\{Y \leq Q_T\})}{f_Y(Q_T)},$

where $\mathbf{1}\{Y \leq Q_T\}$ is an indicator function equal to 1 if the value of the outcome variable is smaller than or equal to the quantile Q_T and 0 otherwise, $f_Y(Q_T)$ is the density of the marginal distribution of Y, and Q_T is the population T-quantile of the *unconditional* distribution of Y. Consequently,

(9) $RIF(Y; Q_T) = Q_T + IF(Y; Q_T).$

In practice, the RIF is first estimated as a function of the sample quantile Q_T (e.g. the 10^{th} percentile), the dummy variable identifying whether the observed outcome, Y, is smaller than or equal to the sample quantile, and the density estimated using kernel methods at the point of the sample quantile. In the second stage, the estimated RIF is used as a dependent variable in an OLS regression that is run separately for the male-managed, female-managed and pooled plot samples. The resulting parameters $\boldsymbol{\gamma}_M$, $\boldsymbol{\gamma}_F$ and $\boldsymbol{\gamma}^*$ replace the $\boldsymbol{\beta}$ counterparts in Equation (6) and are used together with the group-specific mean values for each covariate, X_M and X_F, to perform aggregate and detailed decompositions of any distributional statistic beyond the mean within the framework provided in Section 4.

7 RIF Decomposition Results

7.1 RIF Aggregate Decomposition

Table 6 presents the gender gap estimates and aggregate RIF decompositions at the mean, and at each decile of the agricultural productivity distribution[29]. The graphical representation of these findings are reported in Figure 3. The estimations are underlined by RIF regressions that use the same set of independent variables included in the base specification for the mean decomposition.

Two key findings emerge from Table 6. First, the estimates of the gender gap and the share of the gender gap attributed to female structural disadvantage increase steadily across the agricultural productivity distribution. While the gender gap is estimated at 25.4 and 23.3 percent at the mean and median, respectively, the estimates at the 10th and 90th percentiles are 22.6 and 37.6 percent, respectively. The female structural disadvantage

[29] Table 5 presents the gender gap estimates and aggregate RIF decomposition at the mean and at each decile of the maize yields distribution. The gender gap in maize yields is estimated at 22.4 and 24.1 percent at the mean and median respectively while the estimates at the 10th and 90th percentiles are 22.2 and 20.2 percent respectively. The structural effect accounts for 26.3 percent at the mean and 45.7 percent at the 90th percentile.

component accounts for 17.7 percent of the gender gap at the mean and 33.5 percent of the gender gap at the 90th percentile.

Second, the gender gap at the lowest three deciles of the agricultural productivity distribution is explained fully by differences in observable covariates, with the endowment effect still accounting for close to 90 percent of the gender gap at the median. Given the trends in the female structural disadvantage component, the percentage contribution of the endowment effect toward the gender gap declines throughout and is more of a dominant force in the first half of the distribution.

7.2 RIF Detailed Decomposition

In the interest of brevity, we provide the detailed RIF decompositions at the 10th, 50th and 90th percentiles. The RIF regressions underlying the detailed decompositions are reported in Tables H.8 through H.10 in Appendix H. The key variables that will be the subject of the RIF detailed decomposition discussion are in line with those that have been emphasized as part of the mean decomposition results. While this indicates that the policies need to address these factors for all women, their relative importance in fact changes across the distribution.

Before discussing the detailed decomposition results in depth, we focus on the graphical representations of RIF regression coefficients for key explanatory variables estimated separately within the pooled, male-managed, and female-managed plot samples at each decile of the agricultural productivity distribution.[30] The trends in RIF regression coefficients for (i) the log of household adult male labor, (ii) the share of plot area under export crop cultivation, (iii) the log of inorganic fertilizer use, and (iv) the child dependency ratio are depicted in Figures 4A, 4B, 4C and 4D, respectively. We find that the evolution of the returns to inorganic fertilizer use and the share of plot area under export cultivated area are at odds with one another. The coefficient associated with inorganic fertilizer use declines steadily throughout the agricultural productivity distribution, while the return to the share of plot area under export cultivation is significantly higher at each decile. This result holds true independent of the plot sample in question.

Moreover, the distribution of returns to household adult male labor is considerably different within the male-managed plot sample *vis-à-vis* its female-managed counterpart. The return to household adult male labor on female-managed plots declines steadily and dips below zero starting with the 70th percentile. The coefficient of interest is, conversely, always positive, and displays a stagnant evolution across the deciles within the male-managed plot sample. The evolution of the coefficient associated with household child dependency ratio among male-managed is also always positive across the distribution, but negative at each decile within the female-managed plot sample and is highest, in absolute terms, at the 90th percentile.

[30] The decile-specific RIF regression results for each plot sample are available upon request.

The detailed RIF decompositions at the 10th, 50th, and 90th percentiles are reported in Table 7.[31] We observe that at the median, the decomposition is comparable to its counterpart at the mean. The fact that females manage smaller plots reduces the gender gap through its negative contribution to the endowment effect across the majority of the agricultural productivity distribution. At the first decile, the log of plot area is associated with a 32 percent reduction in the endowment effect. Household adult female labor input is the other key variable that is associated with negative contributions towards both the endowment effect and the male structural advantage component at each decile. Although the magnitude of the relationship between the variable and the endowment effect decreases in relative and absolute terms towards the higher end of the agricultural productivity distribution, it remains economically significant and indicates the importance of household female adult labor in the context of labor market failures and insufficient household male adult labor. The sustained negative contributions towards the male structural advantage components are driven by lower returns to household adult female labor on male-managed plots *vis-à-vis* pooled and female-managed plots.

The log of inorganic fertilizer use per hectare is associated with positive but decreasing contributions towards the endowment effect. The share of plot area cultivated with improved seeds exhibits a similar trend. Addressing gender differences in access to

[31] To lend support towards the assumptions of ignorability and zero conditional mean associated with the RIF decomposition, we follow the added-control approach proposed by Altonji et al. (2005), and implemented in Section 5.3 for the mean decomposition. Tables H.11 through H.19 in the Appendix present the results of the RIF regressions including the additional controls for the pooled, male-managed, and female-managed plot samples for the 10th, 50th and 90th percentiles. We find that the coefficients are largely stable in terms of magnitude, and do not change sign or significance in response to additional control variables that span five domains.

inorganic fertilizer and improved seeds would, therefore, alleviate the gender gap mostly within the first half of the productivity distribution. The sustained increases in returns to the share of plot area under export crop cultivation at each decile of the agricultural productivity distribution underlies, in contrast, the surge in the portion of the endowment effect attributed to this variable. The share of plot area under export crop cultivation accounts for 40 and 56 percent of the gender gap at the 10^{th} and 90^{th} percentile, respectively. Diversification into high-value, export-oriented agriculture among female farmers, independent of productivity level, is, therefore, a clear channel through which large strides could be attained in closing the gender gap.

Furthermore, household adult male labor input at the plot level contributes differently toward the endowment and structure effect at different points of the agricultural productivity distribution. At the 10^{th} percentile, its statistically significant contribution towards the gender gap exists only through the endowment effect, while at the 90^{th} percentile, its effect exists only through the structure effect. The gender gap widening effect of being rationed out of household male labor is more pronounced for female farmers in the first half of the agricultural productivity distribution. At the upper deciles of the agricultural productivity distribution, the variable is associated with higher gender gap instead through its contribution towards the male structural advantage and the female structural disadvantage. While the underlying causes of this pattern need to be studied deeper, informal institutional constraints, including potential supervision difficulties associated with household adult male labor on female-managed plots, that may lead to higher returns to household adult male labor on male-managed plots could be more binding for female farmers of high productivity levels.

Finally, household wealth and access to agricultural implements are associated with sustained, positive contributions towards the endowment effect at each decile of interest. The latter finding marks the importance of access to labor-saving technologies in bridging the gap, especially since the farm duties of female managers are usually compounded by their duties at home. The sustained positive contributions of household child dependency ratio towards the structure effect throughout the productivity distribution lends support to this argument, since, as noted above, the relationship between this variable and agricultural productivity on female-managed plots is consistently negative.

8 Conclusion

This study offers a fresh look at gender differences in sub-Saharan African agricultural productivity, the alleviation of which have been advocated by governments and international donor community as one of the key drivers of broad, agriculture-based economic growth and ensuing gains in living standards. Our contribution to the literature is to (i) apply decomposition techniques that identify the relative quantitative importance of factors explaining the gender gap at the mean and other points of the agricultural productivity distribution, and (ii) to use nationally-representative data, collected within a multi-topic framework and with emphasis on agriculture.

While the gender gap in Malawi is estimated at 25.4 percent at the mean, it ranges from 22.6 percent at the 10th percentile to 37.6 percent at the 90th percentile. The findings support the view that large and significant gender disparities in use of inputs and asset ownership are the central factors behind the gender gap, particularly in the first half of the agricultural productivity distribution. At the mean and the median, the differences in observable covariates are associated with 82 and 87 percent of the gender gap, respectively. Above the median, the percentage contribution of the endowment effect towards the gender gap declines steadily, whereby at the 80th and 90th deciles of the distribution of agricultural productivity, the structure effect, which is driven by gender differences in returns to factors of production, explains 30 and 34 percent of the gender gap, respectively.

Higher levels of household adult male labor and area under export crop cultivation on male-managed plots, in particular, widen the gender gap; a result that holds true across the vast majority of the agricultural productivity distribution. These disparities appear to be compounded by gender differences in the availability of time devoted to productive activities, as negative returns to household child dependency ratio on female managed plots are found to exacerbate the female structural disadvantage component of the gender gap at each decile of the agricultural productivity distribution. In addition, lower and declining returns to household adult male labor on female managed plots *vis-à-vis* male-managed comparators across the agricultural productivity distribution might be suggestive of potential household adult male labor supervision difficulties on female-managed plots. These mutually reinforcing constraints appear to generate a female productivity trap; as such, policies need to prioritize and target the key factors underlying the gender gap.

Our study shows a number of factors that seem to be driving the gender differences in agricultural productivity in Malawi. While we demonstrate that diversification among female farmers into high-value agriculture (with appropriate adoption support and risk mitigation mechanisms), and counteracting the effects of household male labor shortages on female-managed plots with enhanced access to inorganic fertilizer, improved seeds and labor-saving agricultural implements could lead to significant contractions in the gender gap across the agricultural productivity distribution, our analysis alone is not enough to inform effective policy interventions that will ensure the realization of these outcomes. In other words, while we can quantify the relative contributions of various factors towards the gender gap, we cannot determine why inequalities in time use, access and returns to agricultural inputs, and the like persist. Although this limitation is inherent in the use of decomposition methods, our empirical approach identifies the key inequalities that will be the focus of our future research, which will seek to map out their determinants in order to inform policy interventions aimed at addressing the gender gap at its roots in Malawi and other parts of sub-Saharan Africa.

Table 1: Descriptive Statistics & Results from Tests & Mean Differences by Gender of the Plot Manager

	Pooled Sample	*Male-Managed Plot Sample*	*Female-Managed Plot Sample*	*Difference*	
ome Variable					
Gross Value of Output (MK)/HA	53,067	56,810	42,477	14,334	***
Manager Characteristics					
ager & Owner Overlap †	0.63	0.58	0.77	-0.19	***
(Years)	42.97	41.59	46.89	-5.30	***
s of Schooling	5.06	5.67	3.33	2.33	***
tionship to Household Head					
Head †	0.94	0.99	0.80	0.19	***
Wife/Husband †	0.05	0.01	0.19	-0.18	***
Child/Adopted Child †	0.01	0.00	0.01	-0.01	***
Other Relative †	0.00	0.00	0.00	0.00	*
Non-Relative †	0.00	0.00	0.00	0.00	
Area					
-Based Plot Area (HA)	0.39	0.41	0.36	0.05	***
Ownership Status					
usively Male Owned †	0.33	0.43	0.04	0.39	***
usively Female Owned †	0.36	0.23	0.75	-0.53	***
Male-Female Owned †	0.12	0.15	0.03	0.12	***
Owned †	0.19	0.19	0.17	0.01	
Non-Labor Input Use					
lence of Pesticide/Herbicide Use †	0.02	0.02	0.01	0.01	***
lence of Organic Fertilizer Use †	0.12	0.12	0.11	0.01	
lence of Inorganic Fertilizer Use †	0.64	0.65	0.62	0.03	**
ganic Fertilizer Use (KG)/HA conditional]	143.61	147.61	132.29	15.33	***
ganic Fertilizer Use (KG)/HA ditional]	224.35	228.03	213.49	14.54	***
Labor Input Use					
lence of Household Male Labor Use †	0.83	0.97	0.43	0.55	***
sehold Male Labor Use (Hours)/HA [Unconditional]	434.54	526.87	173.32	353.56	***
sehold Male Labor Use (Hours)/HA [Conditional]	523.55	541.51	407.35	134.17	***
lence of Household Female Labor Use †	0.95	0.94	0.98	-0.04	***
sehold Female Labor Use (Hours)/HA [Unconditional]	506.88	455.52	652.18	-196.66	***
sehold Female Labor Use (Hours)/HA [Conditional]	532.83	484.20	664.76	-180.56	***
lence of Household Child Labor Use †	0.25	0.22	0.32	-0.10	***
sehold Child Labor Use (Hours)/HA [Unconditional]	64.59	54.35	93.54	-39.19	***
sehold Child Labor Use (Hours)/HA [Conditional]	261.13	245.86	290.83	-44.98	***
lence of Hired Labor Use †	0.23	0.24	0.23	0.01	
d Labor Use (Days)/HA [Unconditional]	6.46	6.33	6.83	-0.50	
d Labor Use (Days)/HA [Conditional]	27.66	26.80	30.19	-3.38	*
lence of Exchange Labor Use †	0.10	0.09	0.12	-0.04	***
ange Labor Use (Days)/HA [Unconditional]	1.34	1.15	1.87	-0.72	***
ange Labor Use (Days)/HA [Conditional]	14.06	13.52	15.10	-1.58	**
Location					
ation (M)	928.38	946.02	878.46	67.56	***

Table 1 (Cont'd)

	Pooled Sample	*Male-Managed Plot Sample*	*Female-Managed Plot Sample*	*Difference*	
t Location					
tance to Household (KM)	2.17	2.29	1.85	0.44	
t Cultivation					
›rcropped †	0.33	0.30	0.41	-0.11	***
re of Plot Area Under Improved Seeds	0.38	0.39	0.35	0.04	***
re of Plot Area Under Export Crops	0.08	0.10	0.03	0.07	***
ısehold Farm Organization					
nber of Plots Cultivated	2.26	2.34	2.03	0.31	***
tivates…					
1 Plot †	0.27	0.24	0.35	-0.11	***
2 Plots †	0.39	0.38	0.40	-0.02	
3 Plots †	0.22	0.25	0.16	0.08	***
4 Plots †	0.09	0.09	0.07	0.03	***
5+ Plots †	0.04	0.04	0.02	0.02	**
er Household Characteristics					
ısehold Size	4.92	5.14	4.29	0.85	***
ld Dependency Ratio	0.69	0.68	0.71	-0.03	
·icultural Extension Receipt †	0.28	0.29	0.25	0.04	***
:ess to Non-Farm Labor Income †	0.43	0.44	0.39	0.05	***
:ess to Non-Farm Non-Labor Income †	0.22	0.22	0.23	0.00	
alth Index	-0.63	-0.54	-0.89	0.35	***
·icultural Implement Access Index	0.67	0.85	0.16	0.69	***
tance to Nearest ADMARC (KM)	8.03	8.02	8.07	-0.04	
ısehold Agro-Ecological Zone Classification					
pic-warm/semiarid †	0.48	0.49	0.48	0.01	
pic-warm/subhumid †	0.30	0.28	0.35	-0.07	***
pic-cool/semiarid †	0.15	0.16	0.12	0.04	***
pic-cool/subhumid †	0.07	0.07	0.05	0.03	***
ısehold Regional Location					
·th †	0.14	0.15	0.11	0.04	***
ıtral †	0.46	0.49	0.38	0.10	***
ıth †	0.40	0.36	0.51	-0.15	***
servations	16,372	12,029 (73.5%)	4343 (26.5%)		

Note: The estimates are weighted in accordance with the complex survey design. ***/**/* indicate statistical significance at the 1/5/10 percent level, respectively. †denotes a dummy variable.

Table 2: Naïve Regression Results on Gender Differences in Agricultural Productivity

	Dependent Variable: Log[Plot Gross Value of Output (MK)/HA)		
	(1)	(2)	(3)
Female Plot Management †	-0.253***	-0.223***	-0.234***
	(0.023)	(0.023)	(0.022)
Fixed Effects	Agro-Ecological Zones	Regions	Districts
Observations	16,372	16,372	16,372
R-Squared	0.014	0.023	0.065

Note: The estimates are weighted in accordance with the complex survey design.
***/**/* indicate statistical significance at the 1/5/10 percent level, respectively.
†denotes a dummy variable.

Table 3: Base OLS Regression Results Underlying the Mean Decomposition

Dependent Variable: Log[Plot Gross Value of Output (MK)/HA]

	Pooled Sample	*Male-Managed Plot Sample*	*Female-Managed Plot Sample*
Plot Manager Characteristics			
Female †	-0.045*		
	(0.027)		
Manager & Owner Overlap †	0.016	0.020	-0.015
	(0.020)	(0.022)	(0.040)
Age (Years)	-0.001	-0.002*	0.001
	(0.001)	(0.001)	(0.001)
Years of Schooling	0.007**	0.005	0.015***
	(0.003)	(0.003)	(0.005)
Plot Area			
Log[GPS-Based Plot Area (HA)]	-0.282***	-0.261***	-0.296***
	(0.030)	(0.034)	(0.046)
Log[GPS-Based Plot Area (HA) Squared]	0.044***	0.043***	0.042***
	(0.009)	(0.010)	(0.013)
Plot Non-Labor Input Use			
Incidence of Pesticide/Herbicide Use †	0.395***	0.360***	0.491***
	(0.076)	(0.077)	(0.136)
Incidence of Organic Fertilizer Use †	0.043	0.054*	0.017
	(0.027)	(0.032)	(0.045)
Log[Inorganic Fertilizer Use (KG)/HA]	0.077***	0.081***	0.066***
	(0.004)	(0.004)	(0.007)
Plot Labor Input Use			
Log[Household Male Labor Use (Hours)/HA]	0.028***	0.067***	0.005
	(0.005)	(0.009)	(0.007)
Log[Household Female Labor Use (Hours)/HA]	0.032***	0.016**	0.053***
	(0.007)	(0.008)	(0.015)
Log[Household Child Labor Use (Hours)/HA]	0.001	0.006	-0.011*
	(0.004)	(0.005)	(0.006)
Log[Hired Labor Use (Days)/HA]	0.079***	0.080***	0.088***
	(0.008)	(0.009)	(0.014)
Log[Exchange Labor Use (Days)/HA]	0.040***	0.042***	0.033
	(0.012)	(0.014)	(0.020)
Plot Location			
Elevation (M)	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
Distance to Household (KM)	-0.001**	-0.001	-0.002***
	(0.001)	(0.001)	(0.001)
Plot Cultivation			
Intercropped †	0.110***	0.089***	0.165***
	(0.025)	(0.028)	(0.039)
Share of Plot Area Under Improved Seeds	0.099***	0.095***	0.099**
	(0.023)	(0.025)	(0.041)
Share of Plot Area Under Export Crops	1.213***	1.187***	1.255***
	(0.040)	(0.040)	(0.090)

Table 3 (Cont'd)

	Pooled Sample	*Male-Managed Plot Sample*	*Female-Managed Plot Sample*
Household Characteristics			
Household Size	0.014***	0.011*	0.033***
	(0.005)	(0.006)	(0.008)
Child Dependency Ratio	-0.011	0.032	-0.076***
	(0.016)	(0.020)	(0.029)
Agricultural Extension Receipt †	0.077***	0.053**	0.157***
	(0.021)	(0.022)	(0.040)
Access to Non-Farm Labor Income †	-0.076***	-0.075***	-0.057
	(0.019)	(0.021)	(0.037)
Access to Non-Farm Non-Labor Income †	-0.054**	-0.035	-0.097**
	(0.027)	(0.031)	(0.043)
Wealth Index	0.055***	0.062***	0.048***
	(0.006)	(0.006)	(0.011)
Agricultural Implement Access Index	0.042***	0.041***	0.043***
	(0.008)	(0.010)	(0.014)
Distance to Nearest ADMARC (KM)	0.001	0.004	-0.005
	(0.003)	(0.003)	(0.005)
Household Agro-Ecological Zone Classification			
Tropic-warm/semiarid †	0.170**	0.186**	0.071
	(0.071)	(0.077)	(0.085)
Tropic-warm/subhumid †	0.099	0.135*	-0.035
	(0.073)	(0.081)	(0.093)
Tropic-cool/semiarid †	0.073	0.119	-0.103
	(0.079)	(0.086)	(0.090)
Observations	16,372	12,029	4,343
R-Squared	0.336	0.342	0.307

Note: The estimates are weighted in accordance with the complex survey design.
***/**/* indicate statistical significance at the 1/5/10 percent level, respectively. †denotes a dummy variable.

Table 4: Decomposition of the Gender Differential in Agricultural Productivity
Agricultural Productivity Proxied by Log[Plot Gross Value of Output (MK)/HA]

A. Mean Gender Differential			
Mean Male-Managed Plot Agricultural Productivity		10.454***	
		(0.017)	
Mean Female-Managed Plot Agricultural Productivity		10.199***	
		(0.024)	
Mean Gender Differential in Agricultural Productivity		0.254***	
		(0.023)	
B. Aggregate Decomposition	*Endowment Effect*	*Male Structural Advantage*	*Female Structural Disadvantage*
TOTAL	0.209***	0.000	0.045*
	(0.024)	(0.002)	(0.027)
Share of the Gender Differential	82%	0%	18%
C. Detailed Decomposition	*Endowment Effect*	*Male Structural Advantage*	*Female Structural Disadvantage*
Plot Manager Characteristics			
Manager & Owner Overlap †	-0.003	0.002	0.024
	(0.004)	(0.006)	(0.027)
Age (Years)	0.004	-0.036*	-0.072*
	(0.003)	(0.019)	(0.044)
Years of Schooling	0.016**	-0.011	-0.028*
	(0.007)	(0.007)	(0.015)
Plot Area			
Log[GPS-Based Plot Area (HA)]	-0.035***	-0.025	-0.018
	(0.007)	(0.017)	(0.054)
Log[GPS-Based Plot Area (HA) Squared]	-0.016***	-0.001	0.004
	(0.004)	(0.010)	(0.026)
Plot Non-Labor Input Use			
Incidence of Pesticide/Herbicide Use †	0.003***	-0.001	-0.001
	(0.001)	(0.000)	(0.001)
Incidence of Organic Fertilizer Use †	0.000	0.001	0.003
	(0.000)	(0.002)	(0.005)
Log[Inorganic Fertilizer Use (KG)/HA]	0.016***	0.012*	0.035*
	(0.004)	(0.007)	(0.018)
Plot Labor Input Use			
Log[Household Male Labor Use (Hours)/HA]	0.096***	0.220***	0.055***
	(0.018)	(0.043)	(0.011)
Log[Household Female Labor Use (Hours)/HA]	-0.017***	-0.084***	-0.129*
	(0.004)	(0.018)	(0.077)
Log[Household Child Labor Use (Hours)/HA]	-0.000	0.006**	0.020**
	(0.002)	(0.003)	(0.009)
Log[Hired Labor Use (Days)/HA]	0.001	0.001	-0.006
	(0.003)	(0.003)	(0.008)
Log[Exchange Labor Use (Days)/HA]	-0.004***	0.000	0.002
	(0.002)	(0.002)	(0.005)
Plot Location			
Elevation (M)	0.015***	-0.020	-0.035
	(0.004)	(0.021)	(0.054)
Distance to Household (KM)	-0.001	0.001	0.002

Table 4 (Cont'd)

C. Detailed Decomposition (cont.)	*Endowment Effect*	*Male Structural Advantage*	*Female Structural Disadvantage*
Plot Cultivation			
Intercropped †	-0.012***	-0.006	-0.023*
	(0.003)	(0.004)	(0.013)
Share of Plot Area Under Improved Seeds	0.004***	-0.002	-0.000
	(0.001)	(0.005)	(0.012)
Share of Plot Area Under Export Crops	0.084***	-0.003**	-0.001
	(0.007)	(0.001)	(0.002)
Household Characteristics			
Household Size	0.012***	-0.017	-0.080**
	(0.004)	(0.015)	(0.032)
Child Dependency Ratio	0.000	0.029***	0.047***
	(0.001)	(0.011)	(0.016)
Household Characteristics			
Agricultural Extension Receipt †	0.003**	-0.007**	-0.020**
	(0.001)	(0.003)	(0.008)
Access to Non-Farm Labor Income †	-0.004**	0.001	-0.007
	(0.002)	(0.005)	(0.012)
Access to Non-Farm Non-Labor Income †	0.000	0.004	0.010
	(0.001)	(0.003)	(0.009)
Wealth Index	0.019***	-0.004**	-0.007
	(0.004)	(0.002)	(0.008)
Agricultural Implement Access Index	0.029***	-0.000	-0.000
	(0.006)	(0.004)	(0.002)
Distance to Nearest ADMARC (KM)	-0.000	0.024**	0.052**
	(0.000)	(0.010)	(0.022)
Household Agro-Ecological Zone Classification [Aggregated]	-0.003	0.001	0.014
	(0.004)	(0.004)	(0.018)
Observations		16,372	

Note: The estimates are weighted in accordance with the complex survey design. ***/**/* indicate statistical significance at the 1/5/10 percent level, respectively. †denotes a dummy variable.

Table 5: Decomposition of the Gender Differential in Agricultural Productivity At Selected Points of the Agricultural Productivity Distribution

Agricultural Productivty Proxied by Log[Plot Quantity of Maize Production (KG)/HA]

A. Gender Differential	*Mean*	*10th Percentile*	*20th Percentile*	*30th Percentile*	*40th Percentile*	*50th Percentile*	*60th Percentile*	*70th Percentile*	*80th Percentile*	*90th Percentile*
Male-Managed Plot Value	6.959***	5.777***	6.268***	6.569***	6.814***	7.031***	7.236***	7.456***	7.713***	8.040***
	(0.019)	(0.040)	(0.027)	(0.024)	(0.022)	(0.020)	(0.018)	(0.019)	(0.020)	(0.020)
Female-Managed Plot Value	6.734***	5.555***	6.026***	6.322***	6.583***	6.791***	7.006***	7.225***	7.473***	7.837***
	(0.026)	(0.042)	(0.040)	(0.034)	(0.031)	(0.028)	(0.029)	(0.028)	(0.029)	(0.029)
Gender Differential	0.224***	0.222***	0.242***	0.247***	0.231***	0.241***	0.230***	0.231***	0.240***	0.202***
	(0.026)	(0.051)	(0.040)	(0.034)	(0.032)	(0.030)	(0.030)	(0.030)	(0.031)	(0.031)
B. Aggregate Decompostion	*Mean*	*10th Percentile*	*20th Percentile*	*30th Percentile*	*40th Percentile*	*50th Percentile*	*60th Percentile*	*70th Percentile*	*80th Percentile*	*90th Percentile*
Endowment Effect	0.165***	0.250***	0.226***	0.206***	0.183***	0.177***	0.160***	0.130***	0.105***	0.110***
	(0.025)	(0.048)	(0.038)	(0.034)	(0.031)	(0.028)	(0.027)	(0.028)	(0.031)	(0.033)
Share of the Gender Differential	74%	112%	94%	83%	80%	74%	70%	57%	44%	54%
Male Structural Advantage	-0.000	0.000	0.000	0.000	0.000	0.000	-0.000	-0.000	-0.000	-0.000
	(0.001)	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Share of the Gender Differential	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Female Structural Disadvantage	0.059**	-0.028	0.016	0.041	0.047	0.064*	0.069*	0.100***	0.135***	0.093**
	(0.030)	(0.061)	(0.048)	(0.043)	(0.040)	(0.037)	(0.037)	(0.039)	(0.042)	(0.042)
Share of the Gender Differential	26.3%	-12.5%	6.5%	16.8%	20.4%	26.4%	30.2%	43.5%	56.3%	45.7%
Observations	11,763									

Note: The estimates are weighted in accordance with the complex survey design. ***/**/* indicate statistical significance at the 1/5/10 percent level.

Table 6: Aggregate Decomposition of the Gender Differential in Agricultural Productivity At Selected Points of the Agricultural Productivity Distribution

Agricultural Productivty Proxied by Log[Plot Value of Output (MK)/HA]

	Mean	*10th Percentile*	*20th Percentile*	*30th Percentile*	*40th Percentile*	*50th Percentile*	*60th Percentile*	*70th Percentile*	*80th Percentile*	*90th Percentile*
A. Gender Differential										
Male-Managed Plot Value	10.454***	9.202***	9.654***	9.963***	10.220***	10.455***	10.685***	10.932***	11.256***	11.737***
	(0.017)	(0.029)	(0.023)	(0.019)	(0.018)	(0.017)	(0.017)	(0.019)	(0.020)	(0.025)
Female-Managed Plot Value	10.200***	8.976***	9.456***	9.750***	9.997***	10.223***	10.423***	10.671***	10.937***	11.361***
	(0.023)	(0.043)	(0.034)	(0.028)	(0.026)	(0.024)	(0.024)	(0.027)	(0.025)	(0.041)
Gender Differential	0.254***	0.226***	0.198***	0.213***	0.223***	0.233***	0.262***	0.260***	0.319***	0.376***
	(0.023)	(0.045)	(0.036)	(0.029)	(0.027)	(0.025)	(0.024)	(0.028)	(0.028)	(0.043)
B. Aggregate Decomposition										
Endowment Effect	0.209***	0.244***	0.236***	0.218***	0.199***	0.203***	0.188***	0.207***	0.224***	0.250***
	(0.024)	(0.042)	(0.033)	(0.031)	(0.029)	(0.026)	(0.026)	(0.031)	(0.030)	(0.042)
Share of the Gender Differential	82.3%	107.9%	118.8%	102.4%	89.1%	87.3%	71.9%	79.6%	70.2%	66.5%
Male Structural Advantage	0.000	0.000	0.000	-0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.002)	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.003)
Share of the Gender Differential	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Female Structural Disadvantage	0.045*	-0.018	-0.037	-0.005	0.024	0.030	0.074**	0.053	0.095***	0.126**
	(0.027)	(0.053)	(0.044)	(0.037)	(0.037)	(0.034)	(0.033)	(0.037)	(0.035)	(0.052)
Share of the Gender Differential	17.7%	-7.9%	-18.8%	-2.4%	10.9%	12.7%	28.1%	20.4%	29.8%	33.5%
Observations	16,372									

Note: The estimates are weighted in accordance with the complex survey design. ***/**/* indicate statistical significance at the 1/5/10 percent level, respectively.

Table 7: Detailed Decomposition of the Gender Differential in Agricultural Productivity At Selected Points of the Agricultural Productivity Distribution

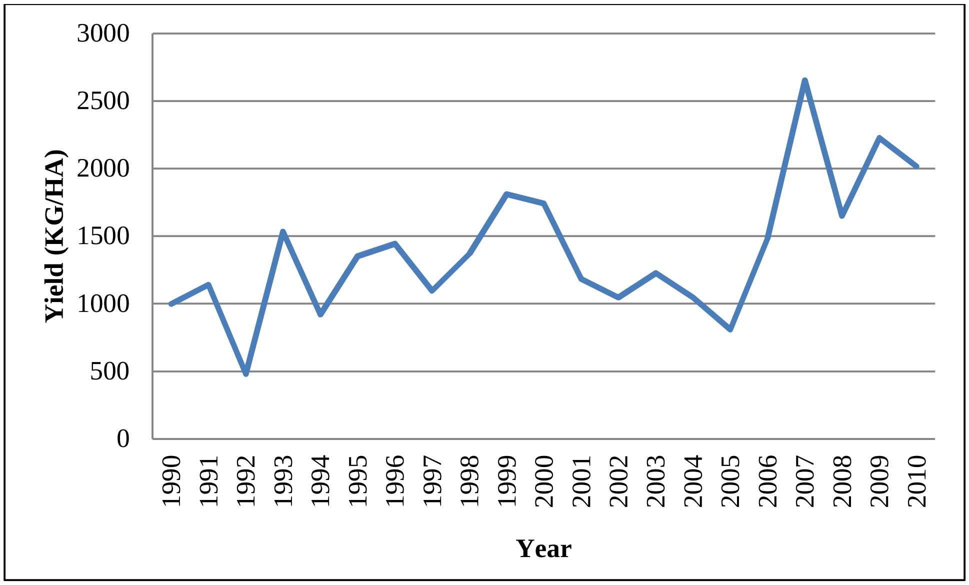
Agricultural Productivty Proxied by Log[Plot Value of Output (MK)/HA]

A. Gender Differential	**Mean**	**10th Percentile**	**50th Percentile**	**90th Percentile**
Male-Managed Plot Value	10.454***	9.202***	10.455***	11.737***
	(0.017)	(0.029)	(0.017)	(0.025)
Female-Managed Plot Value	10.199***	8.976***	10.223***	11.361***
	(0.024)	(0.043)	(0.024)	(0.041)
Gender Differential	0.254***	0.226***	0.233***	0.376***
	(0.023)	(0.045)	(0.025)	(0.043)

B. Aggregate Decomposition	*Endowment Effect*				*Male Structural Advantage*				*Female Structural Disadvantage*			
	Mean	*10th*	*50th*	*90th*	*Mean*	*10th*	*50th*	*90th*	*Mean*	*10th*	*50th*	*90th*
TOTAL	0.209***	0.244***	0.203***	0.250***	0.000	0.000	0.000	0.000	0.045*	-0.018	0.030	0.126**
	(0.024)	(0.042)	(0.026)	(0.042)	(0.002)	(0.003)	(0.002)	(0.003)	(0.027)	(0.053)	(0.034)	(0.052)
Share of the Gender Differential	82.1%	107.9%	87.3%	66.5%	0.0%	0.0%	0.0%	0.0%	17.6%	-7.9%	12.7%	33.5%
C. Detailed Decomposition	*Endowment Effect*				*Male Structural Advantage*				*Female Structural Disadvantage*			
	Mean	*10th*	*50th*	*90th*	*Mean*	*10th*	*50th*	*90th*	*Mean*	*10th*	*50th*	*90th*
Plot Manager Characteristics												
Manager & Owner Overlap †	-0.003	-0.018**	-0.001	0.006	0.002	-0.000	0.003	0.008	0.024	0.006	0.036	0.041
	(0.004)	(0.008)	(0.004)	(0.007)	(0.006)	(0.012)	(0.007)	(0.012)	(0.027)	(0.058)	(0.031)	(0.055)
Age (Years)	0.004	0.005	-0.000	0.006	-0.036*	-0.042	-0.051**	0.017	-0.072*	-0.098	-0.091*	0.021
	(0.003)	(0.008)	(0.004)	(0.006)	(0.019)	(0.048)	(0.021)	(0.037)	(0.044)	(0.106)	(0.049)	(0.087)
Years of Schooling	0.016**	0.022	0.023***	0.007	-0.011	-0.024	-0.007	-0.020	-0.028*	-0.061*	-0.015	-0.053
	(0.007)	(0.014)	(0.008)	(0.011)	(0.007)	(0.016)	(0.009)	(0.017)	(0.015)	(0.032)	(0.018)	(0.037)
Plot Area												
Log[GPS-Based Plot Area (HA)]	-0.035***	-0.078***	-0.036***	0.009	-0.025	-0.054	-0.049***	-0.033	-0.018	-0.161	-0.079	-0.050
	(0.007)	(0.015)	(0.007)	(0.007)	(0.017)	(0.038)	(0.019)	(0.040)	(0.054)	(0.134)	(0.059)	(0.121)
Log[GPS-Based Plot Area (HA) Sq.]	-0.016***	0.023***	-0.011***	-0.071***	-0.001	0.001	0.017*	-0.030	0.004	0.034	0.047*	-0.036
	(0.004)	(0.007)	(0.003)	(0.015)	(0.010)	(0.017)	(0.010)	(0.029)	(0.026)	(0.054)	(0.025)	(0.068)
Plot Non-Labor Input Use												
Incidence of Pesticide/Herbicide Use †	0.003***	0.002*	0.002***	0.005***	-0.001	-0.001*	-0.001**	0.001	-0.001	-0.003	-0.002*	0.005
	(0.001)	(0.001)	(0.001)	(0.002)	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	(0.002)	(0.001)	(0.003)
Incidence of Organic Fertilizer Use †	0.000	0.001	0.001	0.001	0.001	-0.002	0.000	0.005	0.003	-0.005	0.001	0.014
	(0.000)	(0.001)	(0.001)	(0.001)	(0.002)	(0.004)	(0.002)	(0.004)	(0.005)	(0.011)	(0.005)	(0.010)
Log[Inorganic Fertilizer Use (KG)/HA]	0.016***	0.028***	0.015***	0.007***	0.012*	-0.009	0.008	0.009	0.035*	-0.017	0.027	0.032
	(0.004)	(0.008)	(0.004)	(0.002)	(0.007)	(0.016)	(0.009)	(0.014)	(0.018)	(0.042)	(0.022)	(0.037)
Plot Labor Input Use												
Log[HH Male Labor Use (Hours)/HA]	0.096***	0.181***	0.100***	0.046	0.220***	0.128	0.217***	0.404***	0.055***	0.024	0.047***	0.114***
	(0.018)	(0.034)	(0.020)	(0.034)	(0.043)	(0.083)	(0.050)	(0.088)	(0.011)	(0.022)	(0.014)	(0.024)
Log[HH Female Labor Use (Hours)/HA]	-0.017***	-0.026***	-0.013***	-0.015**	-0.084***	-0.115***	-0.080***	-0.128***	-0.129*	-0.372**	-0.099	-0.227
	(0.004)	(0.007)	(0.004)	(0.007)	(0.018)	(0.037)	(0.020)	(0.042)	(0.077)	(0.167)	(0.085)	(0.182)
Log[HH Child Labor Use (Hours)/HA]	-0.000	0.000	-0.000	-0.001	0.006**	0.007	0.006*	0.001	0.020**	0.015	0.022**	0.005
	(0.002)	(0.005)	(0.003)	(0.005)	(0.003)	(0.006)	(0.004)	(0.006)	(0.009)	(0.020)	(0.011)	(0.018)

Table 7 (Cont'd)												
	Endowment Effect				*Male Structural Advantage*				*Female Structural Disadvantage*			
	Mean	*10th*	*50th*	*90th*	*Mean*	*10th*	*50th*	*90th*	*Mean*	*10th*	*50th*	*90th*
Log[Hired Labor Use (Days)/HA]	0.001	0.001	0.001	0.002	0.001	-0.003	0.002	-0.001	-0.006	-0.012	-0.003	-0.017
	(0.003)	(0.002)	(0.003)	(0.004)	(0.003)	(0.006)	(0.003)	(0.007)	(0.008)	(0.015)	(0.010)	(0.019)
Log[Exchange Labor Use (Days)/HA]	-0.004***	-0.005**	-0.005**	-0.002	0.000	0.000	0.001	-0.001	0.002	0.004	0.003	-0.001
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.004)	(0.005)	(0.009)	(0.006)	(0.010)
Plot Location												
Elevation (M)	0.015***	0.035***	0.012***	-0.006	-0.020	-0.032	-0.005	-0.068*	-0.035	-0.081	-0.002	-0.165*
	(0.004)	(0.009)	(0.004)	(0.005)	(0.021)	(0.047)	(0.024)	(0.038)	(0.054)	(0.120)	(0.063)	(0.100)
Distance to Household (KM)	-0.001	-0.001	-0.001	0.001	0.001	-0.000	0.001	0.002	0.002	-0.003	0.002	0.003
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.004)	(0.002)	(0.003)
Plot Cultivation												
Intercropped †	-0.012***	-0.008	-0.014***	-0.013***	-0.006	-0.016*	-0.003	-0.005	-0.023*	-0.051*	-0.012	-0.016
	(0.003)	(0.005)	(0.004)	(0.004)	(0.004)	(0.009)	(0.005)	(0.007)	(0.013)	(0.029)	(0.016)	(0.023)
Share of Plot Area Under Improved Seeds	0.004***	0.006**	0.006***	-0.000	-0.002	0.016	-0.002	-0.002	-0.000	0.048*	-0.002	0.006
	(0.001)	(0.002)	(0.002)	(0.002)	(0.005)	(0.011)	(0.005)	(0.009)	(0.012)	(0.028)	(0.013)	(0.022)
Share of Plot Area Under Export Crops	0.084***	0.019***	0.066***	0.211***	-0.003**	0.004	-0.001	-0.009**	-0.001	0.006	0.003	-0.014
	(0.007)	(0.003)	(0.005)	(0.018)	(0.001)	(0.002)	(0.001)	(0.004)	(0.002)	(0.004)	(0.002)	(0.009)
Household Characteristics												
Household Size	0.012***	0.018**	0.008*	0.013**	-0.017	-0.019	-0.008	-0.055*	-0.080**	-0.089	-0.065*	-0.163***
	(0.004)	(0.008)	(0.005)	(0.007)	(0.015)	(0.031)	(0.018)	(0.029)	(0.032)	(0.064)	(0.038)	(0.063)
Child Dependency Ratio	0.000	-0.001	-0.000	0.001	0.029***	0.041*	0.019	0.038**	0.047***	0.065**	0.031	0.067***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.011)	(0.022)	(0.013)	(0.018)	(0.016)	(0.031)	(0.019)	(0.026)
Household Characteristics												
Agricultural Extension Receipt †	0.003**	0.005*	0.004**	0.003	-0.007**	-0.006	-0.009**	-0.007	-0.020**	-0.015	-0.025**	-0.021
	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.006)	(0.004)	(0.006)	(0.008)	(0.017)	(0.010)	(0.016)
Access to Non-Farm Labor Income †	-0.004**	-0.006**	-0.003**	-0.006**	0.001	0.026**	-0.002	-0.002	-0.007	0.062**	-0.014	-0.024
	(0.002)	(0.003)	(0.002)	(0.003)	(0.005)	(0.011)	(0.005)	(0.009)	(0.012)	(0.027)	(0.014)	(0.025)
Access to Non-Farm Non-Labor Income †	0.000	0.000	0.000	0.000	0.004	0.008	0.008**	0.003	0.010	0.014	0.021**	0.007
	(0.001)	(0.002)	(0.001)	(0.000)	(0.003)	(0.006)	(0.004)	(0.006)	(0.009)	(0.018)	(0.010)	(0.017)
Wealth Index	0.019***	0.018***	0.021***	0.019***	-0.004**	-0.002	-0.005**	-0.001	-0.007	-0.007	-0.014	0.014
	(0.004)	(0.004)	(0.004)	(0.005)	(0.002)	(0.003)	(0.002)	(0.004)	(0.008)	(0.012)	(0.009)	(0.018)
Agricultural Implement Access Index	0.029***	0.031***	0.030***	0.024***	-0.000	-0.011*	-0.001	-0.002	-0.000	-0.007*	-0.000	-0.002
	(0.006)	(0.011)	(0.007)	(0.009)	(0.004)	(0.006)	(0.004)	(0.006)	(0.002)	(0.004)	(0.003)	(0.003)
Distance to Nearest ADMARC (KM)	-0.000	0.000	-0.000	-0.000	0.024**	0.043*	0.016*	0.045***	0.052**	0.105**	0.034	0.101***
	(0.000)	(0.000)	(0.001)	(0.000)	(0.010)	(0.023)	(0.009)	(0.014)	(0.022)	(0.047)	(0.023)	(0.034)
Household Agro-Ecological Zone Classification [Aggregated]	-0.003	-0.008	-0.001	0.002	0.001	-0.006	0.002	-0.005	0.014	0.003	0.015	-0.010
	(0.004)	(0.008)	(0.004)	(0.005)	(0.004)	(0.009)	(0.005)	(0.010)	(0.018)	(0.041)	(0.022)	(0.038)
Observations	16,372				16,372				16,372			

Note: The estimates are weighted in accordance with the complex survey design. ***/**/* indicate statistical significance at the 1/5/10 percent level, respectively.
†denotes a dummy variable



Source: FAOSTAT

Figure 1: Malawi Annual Maize Yield Estimates (1990-2010)

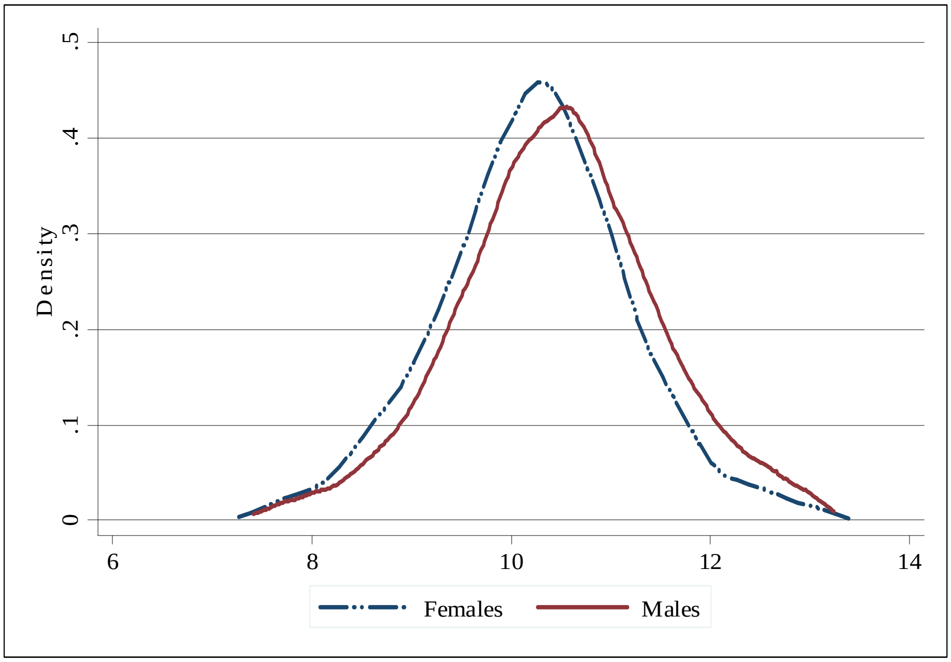


Figure 2: Kernel Density Estimates of the Log of Gross Value of Output per Hectare for Male- and Female-Managed Plot Samples

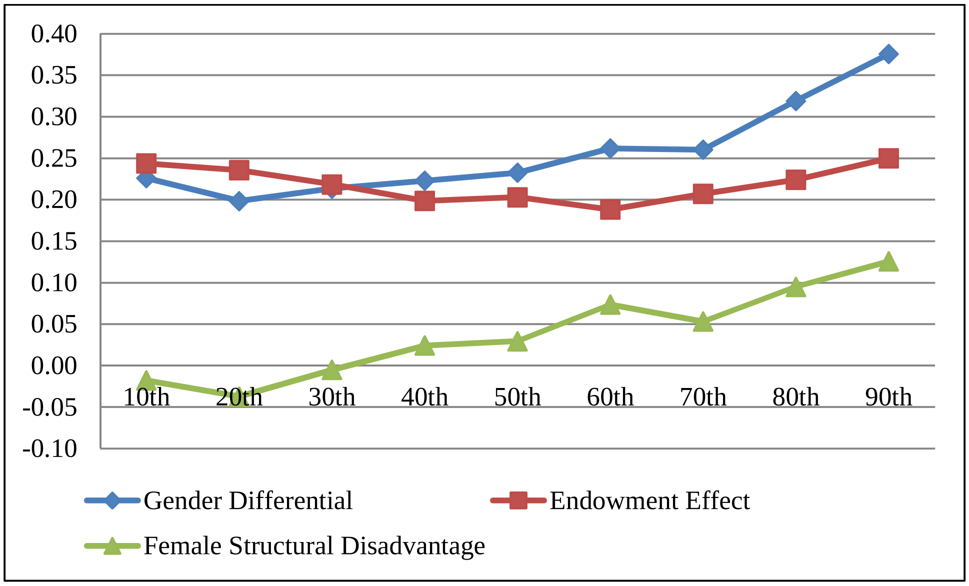
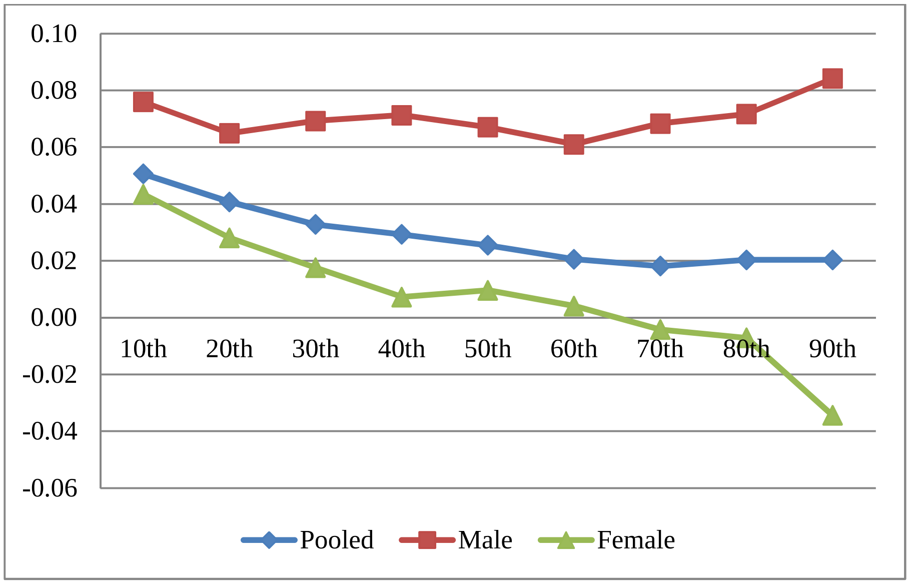
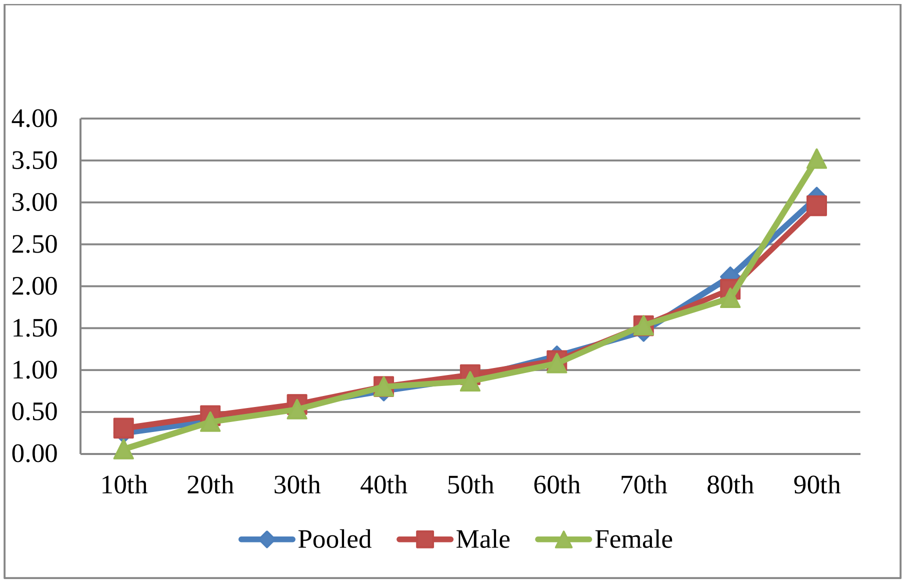


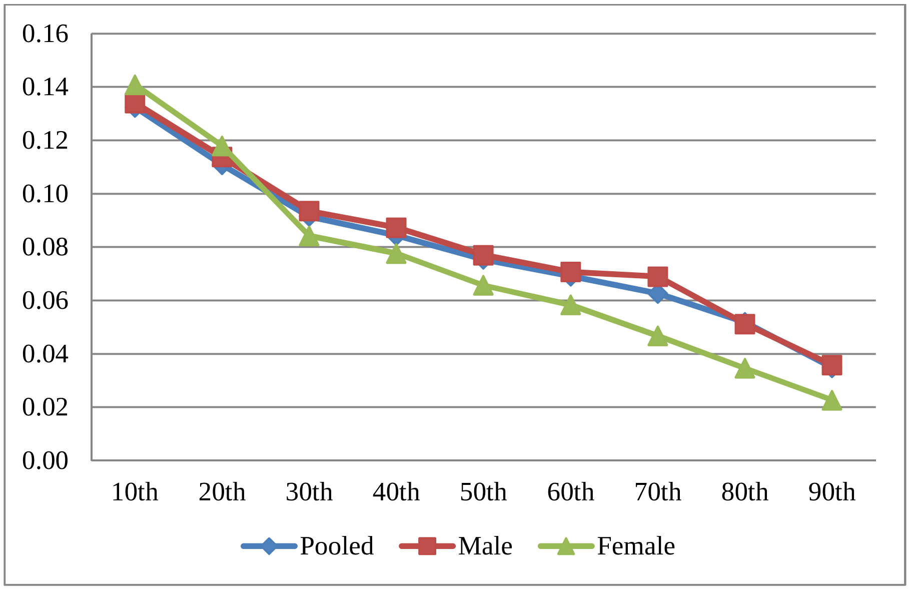
Figure 3: Gender Gap, Endowment Effect and Female Structural Disadvantage Estimated Based on RIF Decomposition at Deciles of Agricultural Productivity Distribution



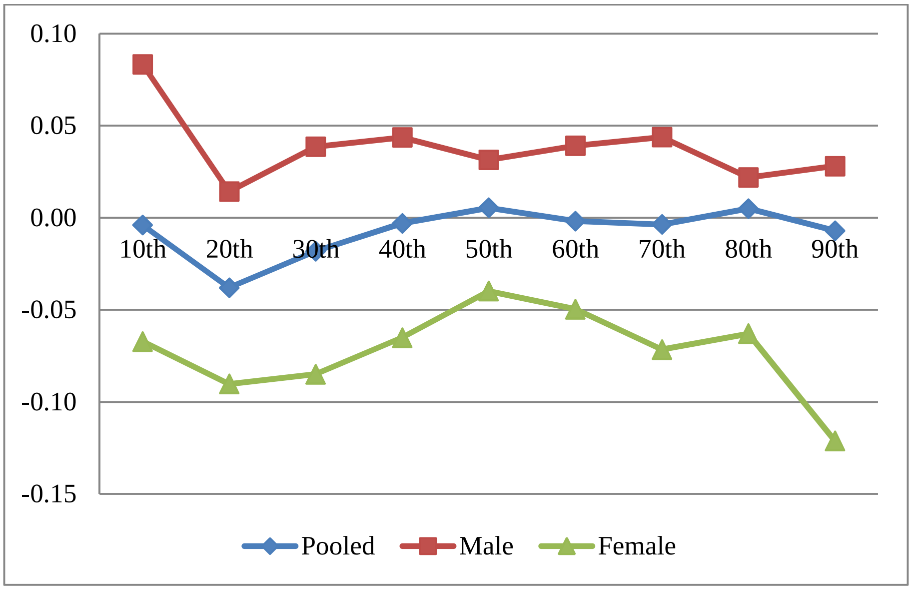
Figures 4A: RIF Regression Coefficients for Plot Log [Household Adult Male Labor Hours/HA], Estimated Separately from Pooled, Male-Managed, and Female-Managed Plot Samples across Deciles of Agricultural Productivity Distribution



Figures 4B: RIF Regression Coefficients for Share of Plot Area Under Export Crop Cultivation, Estimated Separately from Pooled, Male-Managed, and Female-Managed Plot Samples across Deciles of Agricultural Productivity Distribution



Figures 4C: RIF Regression Coefficients for Plot Log[Inorganic Fertilizer Use (KG)/HA], Estimated Separately from Pooled, Male-Managed, and Female-Managed Plot Samples across Deciles of Agricultural Productivity Distribution



Figures 4D: RIF Regression Coefficients for Household Child Dependency Ratio, Estimated Separately from Pooled, Male-Managed, and Female-Managed Plot Samples across Deciles of Agricultural Productivity Distribution

APPENDIX A

GENDER GAP IN LAND PRODUCTIVITY

The *gender gap* "*D*" is expressed as the mean outcome difference:

$$D = E(Y^M) - E(Y^F) \tag{A.1}$$

Equations (12a) and (12b) imply that:

$$E(Y^M) = \beta_0^M + E(L^M)\beta_L^M + \sum_k E(X_k^M)\beta_k^M + \sum_j E(Z_j^M)\delta_j^M \tag{A.2a}$$

$$E(Y^F) = \beta_0^F + E(L^F)\beta_L^F + \sum_k E(X_k^F)\beta_k^F + \sum_j E(Z_j^F)\delta_j^F \tag{A.2b}$$

We rewrite equation (13) using equations (12a) and (12b):

$$D = E(Y^M) - E(Y^F) = \beta_0^M + E(L^M)\beta_L^M + \sum_k E(X_k^M)\beta_k^M + \sum_j E(Z_j^M)\delta_j^M -$$
$$\beta_0^F - E(L^F)\beta_L^F - \sum_k E(X_k^F)\beta_k^F - \sum_j E(Z_j^F)\delta_j^F \tag{A.3}$$

Rearranging Equation (15) by adding and subtracting $E(L^M)\beta_L^F, \sum_k E(X_k^M)\beta_k^F, \sum_j E(Z_j^M)\delta_j^F$, we decompose the gender gap into the following components:

$$D = \underbrace{[E(L^M) - E(L^F)](\beta_L^F)}_{\text{labor market effect}} + \underbrace{\sum_k [E(X_k^M) - E(X_k^F)](\beta_k^F)}_{\text{purchased inputs effect}} + \underbrace{\sum_j [E(Z_j^M) - E(Z_j^F)]\delta_j^F}_{\text{household endowment effect}} +$$

$$\underbrace{E(L^M)[\beta_L^M - (\beta_L^F)] + \sum_k E(X_k^M)[\beta_k^M - \beta_k^F] + (\beta_0^M - \beta_0^F) + \sum_j E(Z_j^M)[\delta_j^M - \delta_j^F]}_{\text{pure marginal productivity effect}}$$

(A.4)

where β_0, β_L, β_k, δ_j are the estimated intercept and slope coefficients of each covariate included in the regressions for the male and female plot samples.

Equation (16) is the aggregate decomposition. The first component is the *labor market effect*, i.e. the portion of the gender gap driven by differences in quantities of labor allocated to farm activities by the head of household. The second component is the *purchased inputs effect*, the portion of the gender gap that is explained by differences in levels of use of inputs that have to be bought such as fertilizer, pesticides, seeds, agricultural implements, and/or hired labor. The third component, the *household endowment effect* is comprised by differences in levels of observable characteristics of the household, including human and physical capital. The fourth component is the *pure marginal productivity effect* and corresponds to the portion of the gender gap explained by differences in the coefficients of each observable covariate included in L and in the *X* and *Z* vectors; as well as differences in the constant between male and female headed households.

APPENDIX B

RECENTERED INFLUENCE FUNCTIONS AND DECOMPOSITION METHODOLOGY

A RIF regression is similar to a standard OLS regression, except that the dependent variable, Y, is replaced by the RIF of the distributional statistic of interest. The approach assumes that the conditional expectation of the $RIF(y;v)$ can be modeled as a linear function of observable attributes, *X*, such that $E[RIF(y;v)\,|\,X] = X\gamma$, as in the mean decomposition. Assuming that $IF(y;v)$ is the influence function corresponding to an observed productivity outcome *y*, for the distributional statistic $v(F_Y)$, the RIF is defined as:

$$RIF(y;v) = v(F_Y) + IF(y;v) \qquad \text{(C.1)}$$

In the case of quantiles, the influence function is equal to:

$$IF(Y;Q_T) = \frac{(T - \mathbf{1}\{Y \leq Q_T\})}{f_Y(Q_T)}, \qquad \text{(C.2)}$$

where $\mathbf{1}\{Y \leq Q_T\}$ is an indicator function equal to 1 if the value of the outcome variable is smaller than or equal to the quantile Q_T and 0 otherwise, $f_Y(Q_T)$ is the density of the marginal distribution of Y, and Q_T is the population T-quantile of the *unconditional* distribution of Y. Consequently,

$$RIF(Y;Q_T) = Q_T + IF(Y;Q_T) \tag{C.3}$$

In practice, the RIF is first estimated as a function of the sample quantile Q_T (e.g. the 10^{th} percentile), the dummy variable identifying whether the observed outcome, Y, is smaller than or equal to the sample quantile, and the density estimated using kernel methods at the point of the sample quantile. In the second stage, the estimated RIF is used as a dependent variable in an OLS regression that is run separately for the male-managed and female-managed samples. The resulting parameters γ^M and γ^F replace the vector of coefficients in Equations (12a) and (12b) and are used together with the group-specific mean values for each covariate to perform aggregate and detailed decompositions of any distributional statistic beyond the mean within the framework provided in Section 3.

APPENDIX C

Table C.1 Scale-Free Normalized Difference

	Male Sample	*Female Sample*	Difference	Normalized Difference
Ln Managerial Labor (hours/ha)	5.81	6.01	-0.20	-0.16
Pesticide/herbicide use yes/no	0.02	0.01	0.01	0.06
Organic fertilizer use yes/no	0.11	0.1	0.01	0.02
Ln Inorganic Fertilizer (kg/ha)	3.32	3.13	0.19	0.05
Ln Hired labor (days/ha)	0.55	0.52	0.03	0.02
Agricultural implements Asset Index	0.84	0.16	0.68	0.38
Proportion of area of the plot under improved varieties	0.38	0.33	0.05	0.08
Proportion of area of the plot under export crops	0.09	0.03	0.06	0.18
Ln GPS Total Area of the plot (ha)	-1.2	-1.29	0.09	0.09
Ln GPS Total Area of the plot (ha) Squared	1.93	2.15	-0.22	-0.08
Elevation (m)	909	843	65.32	0.14
Plot distance to hh	2.06	1.68	0.38	0.03
Inter-croppped	0.32	0.45	-0.13	-0.19
Manager is equal to one of the owners	0.58	0.78	-0.20	-0.31
Age of the manager	41.18	49.07	-7.89	-0.34
Years of Schooling of the manager	5.8	3.21	2.59	0.50
Ln Non-Managerial Household Labor (hours/ha)	5.77	3.8	1.97	0.60
Ln Exchange labor (days/ha)	0.18	0.32	-0.14	-0.13
Household Size	5.16	4	1.16	0.38
Dependency Ratio	0.7	0.71	-0.01	-0.01
Ag extension services receipt	0.32	0.27	0.05	0.08
HH has any off-farm income	0.44	0.35	0.09	0.13
HH receives other transfers/safety net help	0.21	0.23	-0.02	-0.03
Wealth Index	-0.6	-1.03	0.43	0.16
HH Distance (KMs) to Nearest ADMARC	8.19	8.2	-0.01	0.00

Table C.2
List of Additional Controls used in the Added Control Approach

Variable	Data Source
Plot Geospatial Characteristics	
Predominant Soil Type: Sandy †	IHS3
Farmer Assessment of Soil Quality: Good †	IHS3
Irrigated †	IHS3
Percent of Land Classified as Agriculture within 2 Km Radius of Plot Location	GlobCover 2009
Plot Slope (Percentage)	SRTM v4
Potential Wetness Index	AfSIS, TWI
No or Slight Constraint on Nutrient Availability †	HWSD
No or Slight Constraint on Nutrient Retention Capacity †	HWSD
No or Slight Constraint on Rooting Conditions †	HWSD
No or Slight Constraint on Oxygen Availability for Roots †	HWSD
No or Slight Constraint on Excess Salts †	HWSD
No or Slight Constraint on Toxicity †	HWSD
No or Slight Constraint on Workability †	HWSD
Erosion: None †	Soil Map of Malawi
Erosion: Slight †	Soil Map of Malawi
Erosion: Slight to moderate † (omitted category)	Soil Map of Malawi
Soil Depth: Shallow †	Soil Map of Malawi
Soil Depth: Deep †	Soil Map of Malawi
Soil Depth: Very Deep † (omitted category)	Soil Map of Malawi
Surface Texture: Clay, Clay-Loamy †	Soil Map of Malawi
Surface Texture: Loamy, Loamy-Sandy †	Soil Map of Malawi
Surface Texture: Sandy †	Soil Map of Malawi
Surface Texture: Sandy-Clay, Loam-Sandy-Clay †	Soil Map of Malawi
Surface Texture: Sandy-Loam †	Soil Map of Malawi
Surface Texture: Other † (omitted category)	Soil Map of Malawi
Sub-surface Texture: Clay, Clay-Loamy †	Soil Map of Malawi
Sub-surface Texture: Loamy, Loamy-Sandy †	Soil Map of Malawi
Sub-surface Texture: Sandy †	Soil Map of Malawi
Sub-surface: Sandy-Clay, Loam-Sandy-Clay †	Soil Map of Malawi
Sub-surface: Sandy-Loam †	Soil Map of Malawi
Sub-surface: Other † (omitted category)	Soil Map of Malawi
Drainage: Very Poor †	Soil Map of Malawi
Drainage: Poor †	Soil Map of Malawi
Drainage: Poor to Imperfect †	Soil Map of Malawi
Drainage: Imperfect †	Soil Map of Malawi
Drainage: Imperfect to Moderately Well †	Soil Map of Malawi
Drainage: Moderately Well †	Soil Map of Malawi
Drainage: Moderately Well to Well †	Soil Map of Malawi
Drainage: Well †	Soil Map of Malawi
Drainage: Somewhat Excellent † (omitted category)	Soil Map of Malawi
Other Plot Characteristics	
Duration Between Planting and Harvesting (Months)	IHS3
Duration of Last Fallow Period (Years)	IHS3
Tree/Permanent Crops Grown on Plot †	IHS3

Table C.2 (Continued)

Variable	Data Source
Household Characteristics	
# of Household Members 0-5	IHS3
# of Household Members 6-14	IHS3
# of Household Male Members 15-59	IHS3
# of Household Female Members 15-59	IHS3
# of Household Members 60+	IHS3
Ratio of # of Sick Adult Household Members & Total # of Adult Household Members	IHS3
Household Distance to Nearest Road (Eucledian, KMs)	IHS3
Household Distance to Nearest Locality with 20,000+ Population (Eucledian, KMs)	IHS3
Community Characteristics	
Residents Pay Village Headman When Selling or Purchasing Land †	IHS3
Savings and Credit Cooperative in the Community †	IHS3
Distance to Nearest Commercial Bank (KMs)	IHS3
Distance to Nearest Micro-Finance Institution (KMs)	IHS3
Assistant Agriculture Extension Development Officer Lives in the Community †	IHS3
Distance to Nearest Agriculture Extension Development Officer KMs)	IHS3
Irrigation Scheme in the Community	IHS3
# of Fertilizer Sellers in the Community	IHS3
# of Hybrid Maize Seed Sellers in the Community	IHS3
Community Net Receiver of Population †	IHS3

Notes: HWSD: Harmonized World Soil Database
Soil Map of Malawi: Land Resources Evaluation Project
AfSIS TWI: Africa Soil Information Service, Topographic Wetness Index

Table C.3 Added Control Approach for Male Sample

	Base Regression	Enumeration Area Fixed Effects	Plot Geospatial Characteristics	Othe Plot Characteristics	Manager Characteristics	Household Characteristics	Community Characteristics
Labor							
Ln Managerial Labor (hours/ha)	0.227***	0.214***	0.226***	0.227***	0.227***	0.227***	0.224***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Purchased Inputs							
Pesticide/herbicide use yes/no	0.446***	0.430***	0.414***	0.434***	0.446***	0.448***	0.449***
	(0.054)	(0.056)	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)
Organic fertilizer use yes/no	0.021	0.019	0.034	0.022	0.021	0.024	0.020
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
Ln Inorganic Fertilizer Use (kg/ha)	0.070***	0.068***	0.071***	0.070***	0.070***	0.070***	0.071***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Ln Hired labor (days/ha)	0.102***	0.091***	0.099***	0.101***	0.102***	0.102***	0.101***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Agricultural Implements Access Index	0.036***	0.041***	0.035***	0.038***	0.036***	0.036***	0.035***
	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)	(0.007)
Proportion of Plot Area Under Improved Seeds	0.041**	0.032*	0.043**	0.041**	0.041**	0.040**	0.043**
	(0.018)	(0.019)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Proportion of Plot Area Under Export Crops	1.078***	1.066***	1.081***	1.118***	1.078***	1.070***	1.068***
	(0.030)	(0.030)	(0.030)	(0.032)	(0.030)	(0.030)	(0.030)
Household Characteristics and Endowment							
ln GPS Total Area of the plot (ha)	-0.155***	-0.163***	-0.146***	-0.155***	-0.157***	-0.164***	-0.158***
	(0.040)	(0.040)	(0.040)	(0.040)	(0.040)	(0.040)	(0.040)
ln GPS Total Area of the plot (ha) Squared	0.056***	0.053***	0.057***	0.057***	0.056***	0.054***	0.057***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Elevation (m)	0.000**	0.000	0.000	0.000**	0.000**	0.000**	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Plot distance to household	-0.000	-0.000	0.000	-0.000	-0.000	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)

Table C.3 Added Control Approach for Male Sample (continued)

	Base Regression	Enumeration Area Fixed Effects	Plot Geospatial Characteristics	Othe Plot Characteristics	Additional Manager Characteristics	Household Characteristics	Community Characteristics
Intercropped	0.230***	0.264***	0.240***	0.221***	0.231***	0.233***	0.228***
	(0.022)	(0.023)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Manager is equal to one of the owners	-0.007	-0.003	-0.010	-0.008	-0.016	-0.006	-0.003
	(0.016)	(0.017)	(0.016)	(0.016)	(0.017)	(0.016)	(0.016)
Age of the manager	-0.001**	-0.001	-0.001**	-0.001**	-0.002***	0.000	-0.001**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Years of Schooling of the manager	0.003	0.005**	0.002	0.003	0.003	0.003	0.004
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Ln Non-Managerial Household Labor (hours/ha)	0.015**	0.017***	0.016***	0.015**	0.015**	0.017***	0.016**
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Ln Exchange labor (days/ha)	0.042***	0.029**	0.043***	0.041***	0.042***	0.043***	0.043***
	(0.011)	(0.012)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Household Size	0.013***	0.011**	0.014***	0.013***	0.013***	0.001	0.014***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.013)	(0.004)
Dependency Ratio	-0.008	0.005	-0.012	-0.008	-0.008	0.011	-0.009
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.023)	(0.016)
Agricultural Extension Receipt	0.030*	0.031*	0.029*	0.033**	0.028*	0.031*	0.032*
	(0.017)	(0.018)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
HH has any off-farm income	-0.063***	-0.047***	-0.056***	-0.063***	-0.062***	-0.061***	-0.059***
	(0.016)	(0.017)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
HH receives other transfers/safety net help	0.007	-0.009	0.009	0.012	0.008	0.010	0.011
	(0.020)	(0.022)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
Wealth Index	0.062***	0.069***	0.062***	0.061***	0.062***	0.062***	0.063***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Distance to Nearest ADMARC (KM)	0.004**	-0.006	0.004***	0.004**	0.004**	0.003*	0.003
	(0.002)	(0.009)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Number of observations	10,962	10,962	10,710	10,853	10,960	10,962	10,868
R2	0.380	0.462	0.392	0.381	0.380	0.381	0.382
Adjusted R2	0.376	0.423	0.387	0.377	0.377	0.378	0.379

note: *** p<0.01, ** p<0.05, * p<0.1

Table C.4 Added Control Approach for Female Sample

	Base Regression	Enumeration Area Fixed Effects	Plot Geospatial Characteristics	Othe Plot Characteristics	Manager Characteristics	Household Characteristics	Community Characteristics
Labor							
Ln Managerial Labor (hours/ha)	0.234***	0.214***	0.236***	0.235***	0.235***	0.234***	0.231***
	(0.018)	(0.020)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Purchased Inputs							
Pesticide/herbicide use yes/no	0.618***	0.567***	0.642***	0.617***	0.625***	0.614***	0.623***
	(0.121)	(0.131)	(0.125)	(0.121)	(0.121)	(0.121)	(0.121)
Organic fertilizer use yes/no	0.036	0.022	0.034	0.038	0.038	0.041	0.031
	(0.044)	(0.047)	(0.044)	(0.044)	(0.044)	(0.044)	(0.044)
Ln Inorganic Fertilizer Use (kg/ha)	0.058***	0.059***	0.060***	0.057***	0.057***	0.057***	0.058***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Ln Hired labor (days/ha)	0.113***	0.088***	0.113***	0.112***	0.114***	0.114***	0.114***
	(0.012)	(0.014)	(0.013)	(0.012)	(0.012)	(0.012)	(0.012)
Agricultural Implements Access Index	0.048***	0.059***	0.047***	0.046***	0.049***	0.047***	0.046***
	(0.012)	(0.014)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Proportion of Plot Area Under Improved Seeds	0.025	-0.008	0.040	0.025	0.026	0.026	0.026
	(0.034)	(0.036)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)
Proportion of Plot Area Under Export Crops	1.132***	1.115***	1.131***	1.149***	1.130***	1.139***	1.133***
	(0.086)	(0.088)	(0.086)	(0.087)	(0.086)	(0.086)	(0.086)
Household Characteristics and Endowment							
ln GPS Total Area of the plot (ha)	-0.354***	-0.379***	-0.336***	-0.341***	-0.356***	-0.352***	-0.367***
	(0.074)	(0.081)	(0.075)	(0.075)	(0.074)	(0.074)	(0.074)
ln GPS Total Area of the plot (ha) Squared	0.006	0.006	0.013	0.010	0.006	0.007	0.001
	(0.025)	(0.027)	(0.026)	(0.025)	(0.025)	(0.025)	(0.025)
Elevation (m)	0.000**	-0.000	0.000	0.000**	0.000**	0.000**	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Plot distance to household	-0.000	-0.004*	-0.002	0.000	-0.000	-0.000	-0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)

Table C.4 Added Control Approach for Female Sample (continued)

	Base Regression	Enumeration Area Fixed Effects	Plot Geospatial Characteristics	Othe Plot Characteristics	Manager Characteristics	Household Characteristics	Community Characteristics
Intercropped	0.273***	0.308***	0.273***	0.268***	0.275***	0.276***	0.278***
	(0.037)	(0.041)	(0.037)	(0.037)	(0.037)	(0.037)	(0.037)
Manager is equal to one of the owners	-0.002	0.020	0.013	-0.007	-0.019	-0.005	-0.010
	(0.033)	(0.039)	(0.034)	(0.033)	(0.034)	(0.033)	(0.033)
Age of the manager	-0.001	-0.001	-0.001	-0.001	-0.003**	0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Years of Schooling of the manager	0.011**	0.012**	0.011**	0.010**	0.010**	0.010**	0.010**
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Ln Non-Managerial Household Labor (hours/ha)	0.012**	0.003	0.012**	0.012**	0.012**	0.010*	0.012**
	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)	(0.005)
Ln Exchange labor (days/ha)	0.013	0.003	0.014	0.012	0.015	0.015	0.013
	(0.015)	(0.017)	(0.016)	(0.015)	(0.015)	(0.015)	(0.015)
Household Size	0.024***	0.029***	0.025***	0.024***	0.024***	-0.034	0.025***
	(0.008)	(0.010)	(0.008)	(0.008)	(0.008)	(0.040)	(0.008)
Dependency Ratio	-0.068***	-0.076***	-0.074***	-0.069***	-0.069***	-0.049**	-0.067***
	(0.018)	(0.021)	(0.019)	(0.019)	(0.018)	(0.023)	(0.018)
Agricultural Extension Receipt	0.131***	0.127***	0.125***	0.127***	0.127***	0.129***	0.135***
	(0.031)	(0.037)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)
HH has any off-farm income	-0.036	-0.018	-0.019	-0.039	-0.029	-0.035	-0.038
	(0.029)	(0.034)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)
HH receives other transfers/safety net help	-0.030	-0.088**	-0.034	-0.023	-0.031	-0.032	-0.031
	(0.033)	(0.042)	(0.034)	(0.034)	(0.033)	(0.034)	(0.034)
Wealth Index	0.060***	0.073***	0.060***	0.060***	0.063***	0.058***	0.060***
	(0.009)	(0.012)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Distance to Nearest ADMARC (KM)	-0.000	-0.046**	0.003	-0.001	-0.000	0.000	-0.001
	(0.003)	(0.020)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Number of observations	3,242	3,242	3,176	3,207	3,242	3,242	3,225
R2	0.373	0.550	0.386	0.372	0.375	0.375	0.376
Adjusted R2	0.362	0.435	0.368	0.360	0.363	0.363	0.363

note: *** p<0.01, ** p<0.05, * p<0.1

Table C.5 Dropping Extreme Observations

	Male Sample			*Female Sample*		
	Base Regression	Dropping top/bottom 1 percent of Agricultural Land Productivity	Dropping top/bottom 1 percent of Managerial Labor	Base Regression	Dropping top/bottom 1 percent of Agricultural Land Productivity	Dropping top/bottom 1 percent of Managerial Labor
Labor						
Ln Managerial Labor (hours/ha)	0.227***	0.186***	0.208***	0.192***	0.234***	0.201***
	(0.011)	(0.010)	(0.012)	(0.017)	(0.018)	(0.019)
Purchased Inputs						
Pesticide/herbicide use yes/no	0.446***	0.383***	0.429***	0.476***	0.618***	0.690***
	(0.054)	(0.052)	(0.054)	(0.115)	(0.121)	(0.124)
Organic fertilizer use yes/no	0.021	0.025	0.023	0.001	0.036	0.047
	(0.024)	(0.023)	(0.024)	(0.041)	(0.044)	(0.044)
Ln Inorganic Fertilizer Use (kg/ha)	0.070***	0.068***	0.071***	0.055***	0.058***	0.059***
	(0.003)	(0.003)	(0.003)	(0.005)	(0.006)	(0.006)
Ln Hired labor (days/ha)	0.102***	0.093***	0.101***	0.107***	0.113***	0.111***
	(0.007)	(0.007)	(0.007)	(0.011)	(0.012)	(0.013)
Agricultural Implements Access Index	0.036***	0.036***	0.037***	0.032***	0.048***	0.047***
	(0.006)	(0.006)	(0.007)	(0.011)	(0.012)	(0.012)
Proportion of Plot Area Under Improved Seeds	0.041**	0.038**	0.041**	0.024	0.025	0.033
	(0.018)	(0.017)	(0.018)	(0.031)	(0.034)	(0.034)
Proportion of Plot Area Under Export Crops	1.078***	1.010***	1.077***	1.074***	1.132***	1.116***
	(0.030)	(0.029)	(0.030)	(0.081)	(0.086)	(0.087)
Household Characteristics and Endowment						
ln GPS Total Area of the plot (ha)	-0.155***	-0.149***	-0.174***	-0.318***	-0.354***	-0.417***
	(0.040)	(0.038)	(0.041)	(0.070)	(0.074)	(0.075)
ln GPS Total Area of the plot (ha) Squared	0.056***	0.052***	0.049***	0.008	0.006	-0.016
	(0.014)	(0.014)	(0.015)	(0.024)	(0.025)	(0.026)
Elevation (m)	0.000**	0.000***	0.000**	0.000***	0.000**	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Plot distance to household	-0.000	-0.001	-0.000	0.001	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Intercropped	0.230***	0.215***	0.233***	0.265***	0.273***	0.280***
	(0.022)	(0.021)	(0.022)	(0.034)	(0.037)	(0.037)

Table C.5 Dropping Extreme Observations (Continued)

	Male Sample			*Female Sample*		
	Base Regression	Dropping top/bottom 1 percent of Agricultural Land Productivity	Dropping top/bottom 1 percent of Managerial Labor	Base Regression	Dropping top/bottom 1 percent of Agricultural Land Productivity	Dropping top/bottom 1 percent of Managerial Labor
Manager is equal to one of the owners	-0.007	-0.012	-0.012	-0.027	-0.002	0.000
	(0.016)	(0.015)	(0.016)	(0.031)	(0.033)	(0.034)
Age of the manager	-0.001**	-0.001*	-0.001**	-0.000	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Years of Schooling of the manager	0.003	0.004	0.002	0.010**	0.011**	0.011**
	(0.002)	(0.002)	(0.002)	(0.004)	(0.005)	(0.005)
Ln Non-Managerial Household Labor (hours/ha)	0.015**	0.011*	0.014**	0.013**	0.012**	0.012**
	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)
Ln Exchange labor (days/ha)	0.042***	0.042***	0.041***	0.017	0.013	0.014
	(0.011)	(0.011)	(0.011)	(0.014)	(0.015)	(0.015)
Household Size	0.013***	0.015***	0.012***	0.025***	0.024***	0.023***
	(0.004)	(0.004)	(0.004)	(0.007)	(0.008)	(0.008)
Dependency Ratio	-0.008	-0.003	-0.006	-0.061***	-0.068***	-0.065***
	(0.016)	(0.015)	(0.016)	(0.017)	(0.018)	(0.018)
Agricultural Extension Receipt	0.030*	0.026*	0.028*	0.132***	0.131***	0.131***
	(0.017)	(0.016)	(0.017)	(0.029)	(0.031)	(0.031)
HH has any off-farm income	-0.063***	-0.065***	-0.064***	-0.017	-0.036	-0.031
	(0.016)	(0.015)	(0.016)	(0.027)	(0.029)	(0.029)
HH receives other transfers/safety net help	0.007	0.005	0.006	-0.024	-0.030	-0.042
	(0.020)	(0.019)	(0.020)	(0.031)	(0.033)	(0.034)
Wealth Index	0.062***	0.060***	0.064***	0.055***	0.060***	0.059***
	(0.005)	(0.005)	(0.005)	(0.009)	(0.009)	(0.009)
Distance to Nearest ADMARC (KM)	0.004**	0.004**	0.004**	-0.001	-0.000	-0.001
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
Constant	7.958***	8.240***	8.063***	8.544***	8.307***	8.417***
	(0.261)	(0.245)	(0.261)	(0.702)	(0.757)	(0.755)
Number of observations	10,962	10,738	10,755	3,181	3,242	3,164
R2	0.380	0.364	0.368	0.361	0.373	0.357
Adjusted R2	0.376	0.360	0.365	0.349	0.362	0.345

note: *** p<0.01, ** p<0.05, * p<0.1

APPENDIX D

Table D.1
Country List

Austria	Finland	Italy	Spain
Belgium	France	Netherlands	Sweden
Denmark	Germany	Portugal	United Kingdom

Note: Switzerland and Iceland were not included because there is no fiscal government spending and/or energy taxes data available for the period of analysis that uses the same methodology as in the other countries included on the sample. Norway was not included because this country is not an EU country and as such the regulatory framework of the EU may not apply; in addition, it is the world's largest producer of oil and natural gas outside the Middle East (on a per-capita basis), which may set it apart from the other countries.

Table D.2
Description of Variables

Variable	Description	Years Available	Source
Sulfur Dioxide	Year average of daily mean SO2 concentration, micrograms per cubic meter	1995-2006	AirBase from the European Topic Centre on Air and Climate Change, under contract to the European Environment Agency
Nitrogen Dioxide	Year average of daily mean NO2 concentration, micrograms per cubic meter	1995-2008	
Ozone	Year average of daily mean O3 concentration, micrograms per cubic meter	1995-2008	
Household final consumption expenditure per capita (3 year moving average)	Market value of all goods and services including durable products purchased by households.	1989-2008	EUROSTAT
Share of government expenditure on public goods	Government expenditure on public goods over total government expenditure. Including: Public order and safety, Environment protection, Housing and community amenities, Health, Recreation, culture and Religion, Education, Social protection	1989-2008	
Share of total government expenditure over GDP	Total Government Expenditure over GDP	1989-2008	
Trade Intensity	Imports of goods and services plus exports of goods and services over GDP	1994-2008	
Energy Tax Rate	Implicit Tax Rate on Energy	1995-2008	EUROSTAT Statistical Books (2009)
Regulation on Large Utilities (Large Combustion Plant Directive)	The Large Combustion Plant Directive (LCP) refers to the Directive 2001/80/EC of the European Parliament and of the Council of 23 October 2001 on the limitation of emission of certain pollutants into the air from Large Combustion Plants. The LCP Directive entered into force on 27 November 2001. The Directive sets emission level values (ELV's) for SO2, NOX and dust into the air from combustion plants with a rated thermal input equal to or exceeding 50 MW. New combustion plants (licensed after 1 July 1987) must meet the ELVs given in the LCP Directive. A distinction is made between new plants licensed before and after 27 November 2002, with the latter ones having to meet more stringent ELVs. The regulation dummy, takes the value of 1 from 2001 (year in which it was enforced), and 0 otherwise if the country was an EU member in those years.	1990-2008	EEA Report No 2/2007
Regulation on NOx	This variable is reciprocal of the standard emission level values of NOx under the EURO I, II, III, IV Directives for Large Goods Vehicles, measured in g/kWh: 1992-1995=9.0, 1995-1999=7.0, 1999-2005=5.0, 2005-2008=3.5.	1990-2008	EEA Report No 2/2007
Heating Degree Days	Measurement that reflects the demand for energy needed to heat a home or business. The measured used is Relative Degree Days (RDD25) that is the ratio between Actual heating degree-days (ADD) and Mean heating degree-days over period 1980 – 2004 (MDD25) ADD express the severity of the cold in a specific time period taking into consideration outdoor temperature and room temperature. To establish a common and comparable basis, Eurostat defined the following method for the calculation of heating degree days: (18 °C - Tm) x d, if Tm is lower than or equal to 15 °C (heating threshold), where Tm is the mean (Tmin + Tmax / 2) outdoor temperature over a period of d days. Calculations are to be executed on a daily basis (d=1), added up to a calendar month -and subsequently to a year- and published for each Member State separately.	1995-2008	EUROSTAT

Table D.3
Summary Statistics

Variable	Mean	Std. Dev.	Min	Max	Units
SO2	6.73	5.47	0.003	85.47	ug/m3 microgram per Cubic Meter
NO2	29.26	15.72	0.31	120.13	
O3	48.31	13.91	0.96	117.17	
Household final consumption expenditure per capita (3 year moving average)	2.08	1.14	-0.45	5.30	1995 Euros
Share of government expenditure on public goods	0.74	0.03	0.65	0.80	
Share of total government expenditure over GDP	0.47	0.05	0.38	0.63	
Energy Tax Rate	1.66	0.37	0.91	3.16	Euros per Ton of Oil Equivalent
Trade Intensity	0.74	0.26	0.47	1.73	
Heating Degree Days	0.95	0.06	0.80	1.19	

Table D.4
Measures of Variability

	Standard Deviation	Coefficient of Variability
SO2	5.47	0.81
NO2	15.72	0.54
O3	13.91	0.29

Table D.5
Annual Country Averages of SO2 (ug/m3)

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Austria	12.9	9.4	7.5	6.3	5.9	5.2	5.3	5.1	3.8	3.9	3.9	3.0	2.5
Belgium	18.4	15.6	14.1	9.4	8.4	9.2	8.2	8.7	8.4	8.0	8.5	7.6	6.2
Denmark	7.9	3.6	2.7	1.9	2.0	1.0	2.2	2.6	1.7	1.9	1.5	1.5	1.5
Finland	4.1	2.9	2.9	2.0	1.4	1.8	1.9	2.5	2.0	2.0	2.4	2.0	2.0
France					7.6	7.0	6.5	6.8	5.5	5.1	5.1	4.7	
Germany	14.2	9.9	7.4	6.0	5.2	4.8	4.7	5.0	4.2	4.4	4.5	4.1	3.9
Italy				9.9	8.0	7.7	7.9	7.9	6.9	5.3	4.7	4.4	3.9
Netherlands	8.9	6.7	5.6	4.7	4.4	3.7	3.8	3.8	3.5	3.6	3.4	2.5	2.5
Portugal			11.0	10.7	10.6	9.2	7.6	4.6	5.0	4.4	3.2	3.9	2.9
Spain	17.7	13.7	15.9	13.9	10.7	9.6	10.3	8.0	8.2	8.2	7.9	7.7	6.3
Sweden	3.1	2.2	1.6	2.0	1.6	1.8	1.6	1.8	1.7	2.3	1.8	1.6	1.3
United Kingdom	18.1	13.8	12.0	9.5	8.8	9.2	7.2	7.7	5.8	4.7	4.5	4.1	3.8

Table D.6
Annual Country Averages of NO2 (ug/m3)

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Austria	25.2	23.8	23.5	23.1	23.5	22.6	23.0	25.4	24.2	25.3	26.9	24.6	24.1
Belgium	33.9	33.7	29.5	30.9	30.0	32.1	32.1	34.8	32.2	32.4	31.8	30.5	29.5
Denmark	31.5	30.3	24.6	25.4	22.7	22.6	26.3	28.2	25.9	27.2	28.2	24.4	25.4
Finland	24.7	20.8	24.5	16.4	15.0	16.4	18.4	16.8	18.6	15.1	16.6	15.3	14.1
France					31.2	30.4	28.8	31.3	27.5	27.8	26.8	26.3	
Germany	31.7	31.6	29.5	28.1	26.7	26.1	26.3	29.1	26.2	27.6	29.2	28.2	28.5
Italy					46.4	45.8	43.6	45.6	39.9	37.9	38.9	35.2	34.4
Netherlands	35.2	35.3	31.8	31.1	29.5	29.4	29.1	30.7	30.1	28.9	28.4	30.2	31.0
Portugal			28.4	28.2	29.0	29.1	29.9	30.1	28.3	26.8	24.3	26.6	25.0
Spain	60.4	64.9	42.0	41.5	35.9	33.1	26.5	26.1	25.6	26.2	24.2	23.6	20.8
Sweden	8.8	9.0	12.0	14.8	15.9	17.6	18.5	18.4	17.8	17.6	18.5	16.6	17.3
United Kingdom	43.3	43.4	41.1	41.3	38.4	38.7	35.2	38.5	33.4	32.5	32.2	36.1	30.0

Table D.7
Annual Country Averages of O3 (ug/m3)

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Austria	52.1	52.0	55.2	54.7	57.2	57.5	57.5	63.7	55.7	57.5	57.6	55.6	53.7
Belgium	37.2	42.0	43.5	47.1	41.2	42.8	41.6	46.3	41.5	42.1	44.0	39.7	41.4
Denmark			54.7	51.3		51.5	48.8	51.1	49.9	47.5	52.3	50.8	50.9
Finland	61.3	61.4	57.7	61.2	55.6	56.7	60.6	57.3	56.4	58.8	58.5	51.6	52.8
France					45.2	47.9	49.1	54.4	49.0	50.3	51.1	48.3	
Germany	40.9	42.6	44.9	46.9	43.8	45.0	45.4	51.8	47.3	47.4	49.6	47.1	47.0
Italy					48.1	48.6	44.3	54.2	50.0	51.3	54.2	51.2	49.0
Netherlands	34.4	34.7	36.2	40.1	35.3	37.2	37.1	40.6	40.1	38.0	40.1	38.8	38.0
Portugal			28.8	27.4	36.9	41.3	42.6	47.6	47.3	51.8	52.6	53.2	50.8
Spain			46.6	45.7	46.2	47.7	46.6	49.7	47.2	49.8	50.2	49.9	51.6
Sweden	59.2	56.9	52.6	57.0	54.2	53.2	59.9	56.9	57.9	54.9	58.2	54.7	55.9
United Kingdom	40.0	36.7	38.9	42.4	40.5	39.6	41.0	44.3	44.1	43.4	47.2	43.7	47.9

Figure D.1
Share of Government Expenditure on Public Goods (1995-2008)

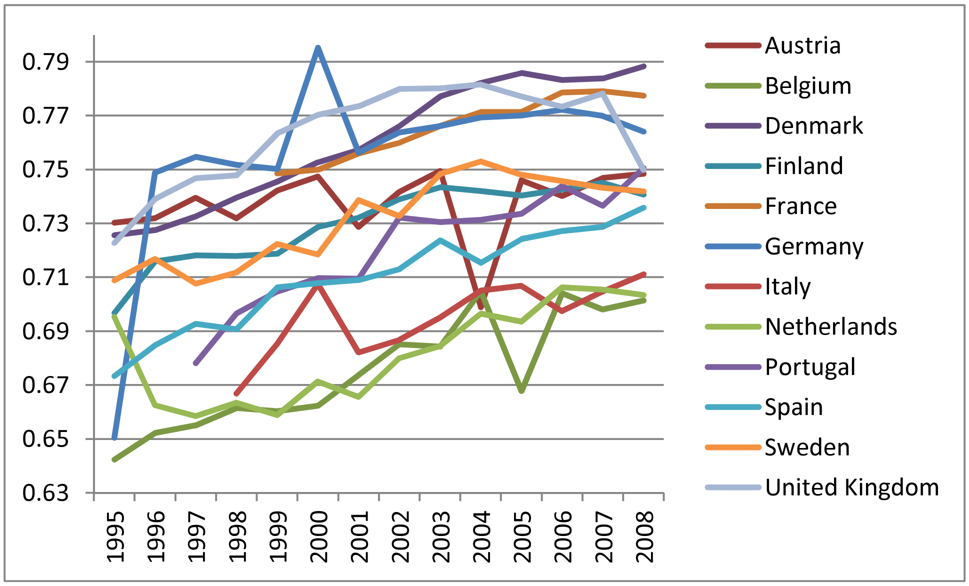
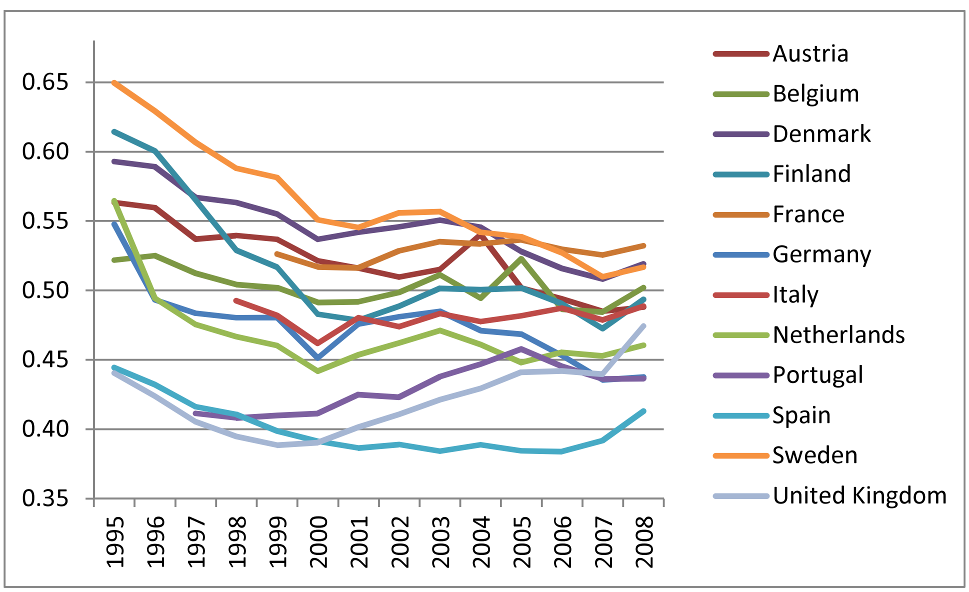


Figure D.2
Share of Total Government Expenditure over GDP (1995-2008)



TableE.1
Comparison of Different Specifications of the Regressions of SO2

	RSE	*RSE with country Fixed Effects*	*RSE with Country Fixed Effects and Year Fixed Effects*	*RSE with TVCE order 1*	*RSE with TVCE order 2*	*RSE with TVCE order 3*	*RSE with TVCE order 4*	*RSE with TVCE order 5*	*RSE with TVCE order 6*
es in public goods over total	-0.027	-0.58**	-0.06	-0.41	-1.18**	-1.08	-2.52**	-3.60**	-5.32**
itures (lagged)	[0.10]	[0.18]	[0.18]	[0.22]	[0.39]	[0.56]	[0.79]	[0.92]	[1.27]
nment expenditures over	0.20**	-0.89**	-0.2	0.48	-0.95	-0.53	-2.19	-3.12*	-5.52**
	[0.07]	[0.24]	[0.27]	[0.40]	[0.64]	[0.83]	[1.18]	[1.32]	[1.78]
Energy Tax Rate	-0.01	-0.04	0.01	0.05	0.03	0.01	0.01	-0.04	-0.11
	[0.03]	[0.03]	[0.03]	[0.04]	[0.04]	[0.05]	[0.05]	[0.05]	[0.06]
Regulation over large Plants	-0.035**	-0.06**	-0.49**	-0.46**	-0.45**	-0.46**	-0.46**	-0.48**	-0.49**
	[0.01]	[0.01]	[0.06]	[0.06]	[0.06]	[0.06]	[0.06]	[0.06]	[0.07]
Log of Trade (X+M)/GDP	-0.95**	-1.06**	-0.14	-0.56*	-0.59*	-0.59*	-0.70*	-0.73*	-1.13**
	[0.10]	[0.12]	[0.20]	[0.22]	[0.23]	[0.26]	[0.28]	[0.29]	[0.40]
3-Year Moving Average of	0.01**	0.02**	0.03**	0.05**	0.05**	0.05**	0.05**	0.04*	0.05**
al consumption per capita	[0.003]	[0.006]	[0.007]	[0.008]	[0.009]	[0.01]	[0.01]	[0.02]	[0.02]

%; ** significant at 1%

Table E.2
Comparison of Different Specifications of the Regressions of NO2

	RSE	*RSE with country Fixed Effects*	*RSE with Country Fixed Effects and Year Fixed Effects*	*RSE with TVCE order 1*	*RSE with TVCE order 2*	*RSE with TVCE order 3*	*RSE with TVCE order 4*	*RSE with TVCE order 5*	*RSE with TVCE order 6*
Share of expenditures in public goods over total government expenditures (lagged)	0.01 [0.05]	-0.12 [0.08]	0.29** [0.09]	0.20 [0.11]	-0.17 [0.17]	-0.10 [0.24]	-0.33 [0.30]	-0.30 [0.35]	-0.19 [0.49]
Share of total government expenditures over GDP (lagged)	0.11** [0.04]	-0.06 [0.10]	-0.15 [0.12]	-0.23 [0.19]	-0.90** [0.28]	-0.41 [0.34]	-0.67 [0.45]	-0.70 [0.51]	-0.37 [0.71]
Time difference of Energy Tax Rate	-0.07** [0.02]	-0.08** [0.02]	-0.14** [0.02]	-0.13** [0.02]	-0.13** [0.0221]	-0.15** [0.02]	-0.18** [0.03]	-0.19** [0.03]	-0.18** [0.03]
Time difference of Regulation over NOx	0.10* [0.05]	0.12* [0.05]	-0.81 [0.52]	-0.84 [0.52]	-0.76 [0.51]	-0.84 [0.51]	-0.90 [0.51]	-0.83 [0.52]	-0.34 [0.51]
Time difference of Log of Trade (X+M)/GDP	-0.21** [0.05]	-0.26** [0.06]	0.21* [0.11]	0.26* [0.12]	0.24 [0.13]	0.12 [0.14]	0.179 [0.15]	0.1 [0.17]	-0.41 [0.25]
Time difference of 3-Year Moving Average of Ln of Household final consumption per capita	0.003 [0.002]	0.01 [0.003]	0.01 [0.005]	0.01 [0.005]	0.02** [0.005]	0.004 [0.007]	0.001 [0.009]	0.005 [0.01]	0.004 [0.01]

* significant at 5%; ** significant at 1%

Table E.3
Comparison of Different Specifications of the Regressions of O3

	RSE	*RSE with country Fixed Effects*	*RSE with Country Fixed Effects and Year Fixed Effects*	*RSE with TVCE order 1*	*RSE with TVCE order 2*	*RSE with TVCE order 3*	*RSE with TVCE order 4*	*RSE with TVCE order 5*	*RSE with TVCE order 6*
Share of expenditures in public goods over total government expenditures (lagged)	-0.0298 [0.08]	-0.40** [0.08]	-0.26** [0.09]	-0.25* [0.11]	-0.24 [0.178]	-0.58* [0.26]	-0.634 [0.33]	-1.21** [0.40]	-1.70** [0.52]
Share of total government expenditures over GDP (lagged)	-0.0655 [0.06]	-0.47** [0.10]	-0.48** [0.13]	-0.31 [0.18]	-0.34 [0.28]	-0.99** [0.37]	-0.10* [0.45]	-1.65** [0.52]	-1.73* [0.77]
Time difference of Energy Tax Rate	0.10** [0.02]	0.09** [0.02]	0.03 [0.02]	0.05* [0.02]	0.03 [0.03]	0.04 [0.03]	0.04 [0.03]	0.07* [0.03]	0.06 [0.04]
Time difference of Regulation over NOx	0.17** [0.04]	0.18** [0.04]	1.44 [0.86]	1.60 [0.86]	1.49 [0.86]	1.37 [0.86]	1.35 [0.86]	1.47 [0.86]	1.54 [0.83]
Time difference of Log of Trade (X+M)/GDP	-0.44** [0.07]	-0.44** [0.07]	-0.04 [0.15]	-0.17 [0.16]	-0.09 [0.17]	-0.04 [0.19]	0.02 [0.20]	-0.02 [0.23]	-0.22 [0.43]
Time difference of 3-Year Moving Average of Ln of Household final consumption per capita	0.002 [0.003]	0.001 [0.003]	-0.006 [0.005]	-0.001 [0.005]	0.004 [0.005]	0.02* [0.008]	0.03** [0.01]	0.03** [0.01]	0.03* [0.01]

* significant at 5%; ** significant at 1%

APPENDIX F

F.1 Extreme Observations Checks

Dropping the top and bottom 1% of the observations on each year.

Table F.1
Coefficient of Share of Expenditures in Public Goods in the RSE-TVCE Regressions

Regression	Bottom 1% of Share of Public Goods Expenditures	Top 1% of Share of Public Goods Expenditures	Top and Bottom 1% of Share of Public Goods Expenditures
SO2	-5.35**	-5.33**	-4.33**
O3	-1.40**	-1.65**	-3.13**

* significant at 5%; ** significant at 1%

Table F.2
Coefficient of Share of Expenditures in Public Goods in the RSE-TVCE Regressions

Regression	Bottom 1% of Pollutant	Top 1% of Pollutant	Top and Bottom 1% of Pollutant
SO2	-6.34**	-4.22**	-5.04**
O3	-1.66**	-1.13*	-1.19**

* significant at 5%; ** significant at 1%

Table F.3
Coefficient of Energy Tax in the RSE-TVCE Regressions

Regressions	Bottom 1% of Energy Tax Rate	Top 1% of Energy Tax Rate	Top and Bottom 1% of Energy Tax Rate
NO2	-0.19**	-0.19**	-0.32**

* significant at 5%; ** significant at 1%

Table F.4
Coefficient of Energy Tax in the RSE-TVCE Regressions

Regressions	Bottom 1% of Pollutant	Top 1% of Pollutant	Top and Bottom 1% of Pollutant
NO2	-0.19**	-0.19**	-0.20**

* significant at 5%; ** significant at 1%

Table F.5
Coefficient of Trade Intensity in the RSE-TVCE Regressions

Regressions	Bottom 1% of Trade Intensity	Top 1% of Trade Intensity	Top and Bottom 1% of Trade Intensity
SO2	-1.19**	-0.93*	-1.04*

* significant at 5%; ** significant at 1%

Table F.6
Coefficient of Trade Intensity in the RSE-TVCE Regressions

Regressions	Bottom 1% of Pollutant	Top 1% of Pollutant	Top and Bottom 1% of Pollutant
SO2	-0.21	-1.27**	-0.46

* significant at 5%; ** significant at 1%

F.2 Country dominance Checks
Dropping one country in each estimation

Figure F.1
Coefficient of the Share of public Goods in the RSE-TVCE Regression for SO2

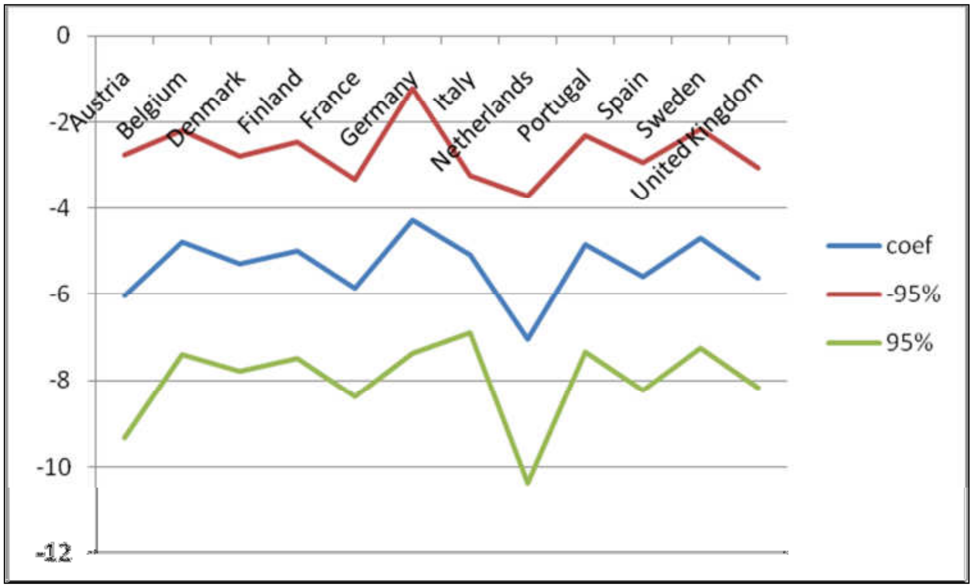


Figure F.2
Coefficient of the Share of public Goods in the RSE-TVCE Regression for O3

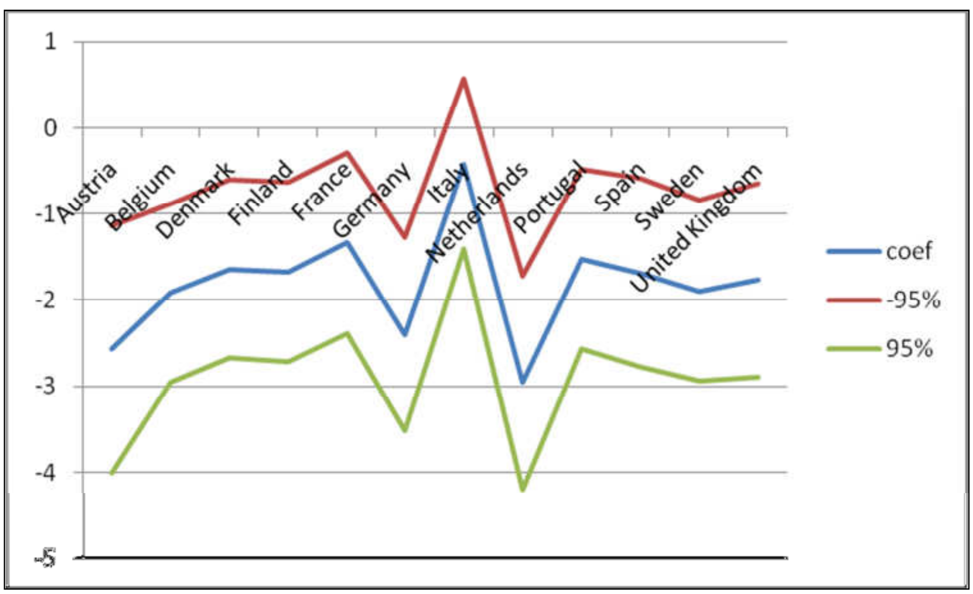


Figure F.3
Coefficient of the Energy Taxes in the RSE-TVCE Regression for NO2

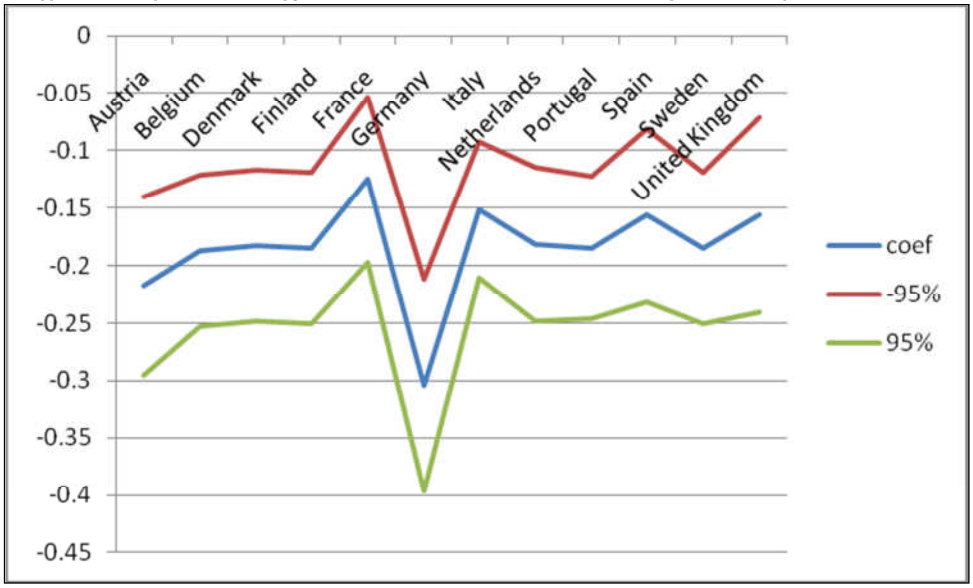
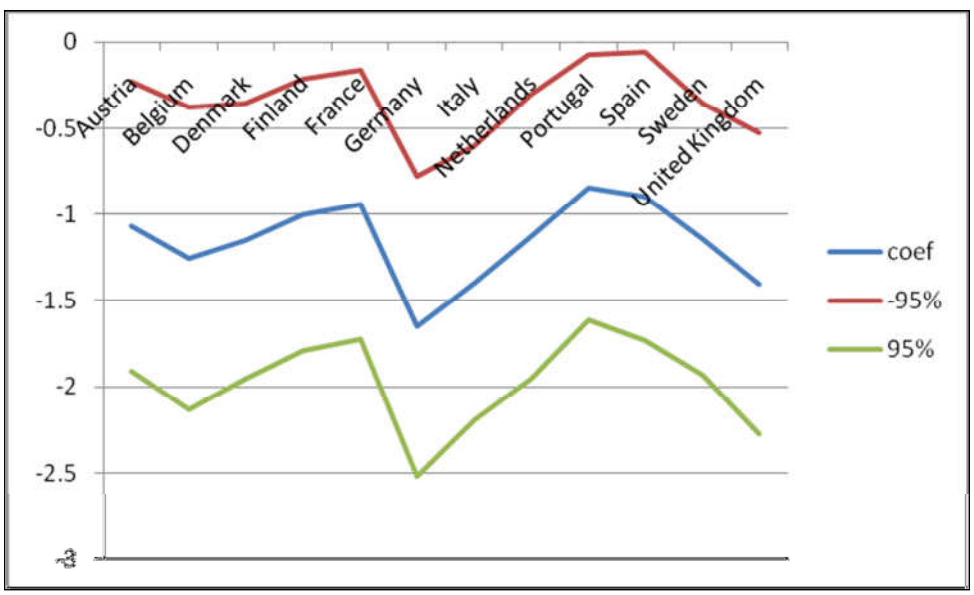


Figure F.4
Coefficient of the Trade Intensity in the RSE-TVCE Regression for SO2



APPENDIX G

Table G.1
Analysis of the Predicted Values of the Time-Varying Country Effects (ν_{jt})

	SO2	NO2	O3
Number of countries with $b_{1j}=b_{2j}=b_{3j}=0$	4	5	5
Signs of the Predicted ν_{jt} Values			
Number of countries with positive predicted values for all years	0	0	0
Number of countries with negative predicted values for all years	6	0	4
Number of countries with predicted values that change sign over time	5	11	7
Monotonicity of the Predicted ν_{jt} Values			
Number of countries with monotonic predicted values over time	0	0	0
Number of countries with one turning point in the predicted values	3	0	0
Number of countries with two turning points in the predicted values	8	11	11

APPENDIX H

Table H.1 Scale-Free Normalized Difference

	Male-Managed Plot Sample	*Female-Managed Plot Sample*	*Difference*	*Normalized Difference*
Plot Manager Characteristics				
Manager & Owner Overlap †	0.58	0.77	-0.19	-0.19
Age (Years)	41.59	46.89	-5.30	-0.14
Years of Schooling	5.67	3.33	2.34	0.28
Plot Area				
Log[GPS-Based Plot Area (HA)]	-1.18	-1.31	0.13	0.07
Log[GPS-Based Plot Area (HA) Squared]	2.00	2.36	-0.36	-0.06
Plot Non-Labor Input Use				
Incidence of Pesticide/Herbicide Use †	0.02	0.01	0.01	0.04
Incidence of Organic Fertilizer Use †	0.12	0.11	0.01	0.01
Log[Inorganic Fertilizer Use (KG)/HA]	3.36	3.16	0.20	0.03
Plot Labor Input Use				
Log[Household Male Labor Use (Hours)/HA]	5.76	2.38	3.38	0.59
Log[Household Female Labor Use (Hours)/HA]	5.47	6.00	-0.53	-0.17
Log[Household Child Labor Use (Hours)/HA]	1.08	1.65	-0.57	-0.10
Log[Hired Labor Use (Days)/HA]	0.67	0.65	0.02	0.01
Log[Exchange Labor Use (Days)/HA]	0.20	0.31	-0.11	-0.06
Plot Location				
Elevation (M)	946.02	878.46	67.56	0.10
Distance to Household (KM)	2.29	1.85	0.44	0.02
Plot Cultivation				
Intercropped †	0.30	0.41	-0.11	-0.10
Share of Plot Area Under Improved Seeds	0.39	0.35	0.04	0.04
Share of Plot Area Under Export Crops	0.10	0.03	0.07	0.15
Household Characteristics				
Household Size	5.14	4.29	0.85	0.17
Child Dependency Ratio	0.68	0.71	-0.03	-0.02
Agricultural Extension Receipt †				
Household Characteristics	0.29	0.25	0.04	0.04
Access to Non-Farm Labor Income †	0.44	0.39	0.05	0.04
Access to Non-Farm Non-Labor Income †	0.22	0.23	-0.01	-0.01
Wealth Index	-0.54	-0.89	0.35	0.08
Agricultural Implement Access Index	0.85	0.16	0.69	0.23
Distance to Nearest ADMARC (KM)	8.02	8.07	-0.05	0.00
Household Agro-Ecological Zone Classification				
Tropic-warm/semiarid †	0.49	0.48	0.01	0.01
Tropic-warm/subhumid †	0.28	0.35	-0.07	-0.07
Tropic-cool/semiarid †	0.16	0.12	0.04	0.05

Table H.2: List of Additional Controls

Variable	Data Source
Plot Geospatial Characteristics	
Predominant Soil Type: Sandy †	IHS3
Farmer Assessment of Soil Quality: Good †	IHS3
Irrigated †	IHS3
Percent of Land Classified as Agriculture within 2 Km Radius of Plot Location	GlobCover 2009
Plot Slope (Percentage)	SRTM v4
Potential Wetness Index	Africa Soil Information Service (AfSIS): Topographic
No or Slight Constraint on Nutrient Availability †	Harmonized World Soil Database (HWSD)
No or Slight Constraint on Nutrient Retention Capacity †	Harmonized World Soil Database (HWSD)
No or Slight Constraint on Rooting Conditions †	Harmonized World Soil Database (HWSD)
No or Slight Constraint on Oxygen Availability for Roots †	Harmonized World Soil Database (HWSD)
No or Slight Constraint on Excess Salts †	Harmonized World Soil Database (HWSD)
No or Slight Constraint on Toxicity †	Harmonized World Soil Database (HWSD)
No or Slight Constraint on Workability †	Harmonized World Soil Database (HWSD)
Erosion: None †	
Erosion: Slight †	
Erosion: Slight to moderate † (omitted category)	
Soil Depth: Shallow †	
Soil Depth: Deep †	
Soil Depth: Very Deep † (omitted category)	
Surface Texture: Clay, Clay-Loamy †	
Surface Texture: Loamy, Loamy-Sandy †	
Surface Texture: Sandy †	
Surface Texture: Sandy-Clay, Loam-Sandy-Clay †	
Surface Texture: Sandy-Loam †	
Surface Texture: Other † (omitted category)	
Sub-surface Texture: Clay, Clay-Loamy †	
Sub-surface Texture: Loamy, Loamy-Sandy †	Soil Map of Malawi, Land Resources Evaluation Project
Sub-surface Texture: Sandy †	
Sub-surface: Sandy-Clay, Loam-Sandy-Clay †	
Sub-surface: Sandy-Loam †	
Sub-surface: Other † (omitted category)	
Drainage: Very Poor †	
Drainage: Poor †	
Drainage: Poor to Imperfect †	
Drainage: Imperfect †	
Drainage: Imperfect to Moderately Well †	
Drainage: Moderately Well †	
Drainage: Moderately Well to Well †	
Drainage: Well †	
Drainage: Somewhat Excellent † (omitted category)	
Other Plot Characteristics	
Duration Between Planting and Harvesting (Months)	IHS3
Duration of Last Fallow Period (Years)	IHS3
Tree/Permanent Crops Grown on Plot †	IHS3

Table H.2 (Cont'd)

Variable	Data Source
Household Characteristics	
# of Household Members 0-5	IHS3
# of Household Members 6-14	IHS3
# of Household Male Members 15-59	IHS3
# of Household Female Members 15-59	IHS3
# of Household Members 60+	IHS3
Ratio of # of Sick Adult Household Members & Total # of Adult Household Members	IHS3
Household Distance to Nearest Road (Eucledian, KMs)	IHS3
Household Distance to Nearest Locality with 20,000+ Population (Eucledian, KMs)	IHS3
Community Characteristics	
Residents Pay Village Headman When Selling or Purchasing Land †	IHS3
Savings and Credit Cooperative in the Community †	IHS3
Distance to Nearest Commercial Bank (KMs)	IHS3
Distance to Nearest Micro-Finance Institution (KMs)	IHS3
Assistant Agriculture Extension Development Officer Lives in the Community †	IHS3
Distance to Nearest Agriculture Extension Development Officer KMs)	IHS3
Irrigation Scheme in the Community	IHS3
# of Fertilizer Sellers in the Community	IHS3
# of Hybrid Maize Seed Sellers in the Community	IHS3
Community Net Receiver of Population †	IHS3

Table H.3: Exploring the Presence of Omitted Variable Bias in Base OLS Regression Results Underlying the Mean Decomposition

Dependent Variable: Log[Plot Value of Output (MK)/HA]

Pooled Sample

		Category of Additional Covariates Integrated into the Base Regression				
	Base	*District Fixed Effects*	*Plot Geospatial Characteristics*	*Other Plot Characteristics*	*Household Characteristics*	*Community Characteristics*
Plot Manager Characteristics						
Female †	-0.045*	-0.059**	-0.051*	-0.049*	-0.048*	-0.041
	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
Manager & Owner Overlap †	0.016	0.004	0.022	0.015	0.016	0.014
	(0.020)	(0.019)	(0.019)	(0.020)	(0.020)	(0.020)
Age (Years)	-0.001	-0.001	-0.001	-0.001	0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Years of Schooling	0.007**	0.003	0.006**	0.007**	0.007**	0.007**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Plot Area						
Log[GPS-Based Plot Area (HA)]	-0.282***	-0.296***	-0.284***	-0.284***	-0.278***	-0.287***
	(0.030)	(0.029)	(0.030)	(0.030)	(0.030)	(0.029)
Log[GPS-Based Plot Area (HA) Squared]	0.044***	0.045***	0.045***	0.044***	0.045***	0.043***
	(0.009)	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)
Plot Non-Labor Input Use						
Incidence of Pesticide/Herbicide Use †	0.395***	0.436***	0.377***	0.390***	0.397***	0.397***
	(0.076)	(0.068)	(0.074)	(0.078)	(0.077)	(0.075)
Incidence of Organic Fertilizer Use †	0.043	0.042	0.051**	0.046*	0.043	0.041
	(0.027)	(0.027)	(0.026)	(0.028)	(0.027)	(0.027)
Log[Inorganic Fertilizer Use (KG)/HA]	0.077***	0.081***	0.080***	0.077***	0.077***	0.078***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Plot Labor Input Use						
Log[Household Male Labor Use (Hours)/HA]	0.028***	0.027***	0.027***	0.028***	0.030***	0.029***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)
Log[Household Female Labor Use (Hours)/HA]	0.032***	0.030***	0.031***	0.032***	0.034***	0.032***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)	(0.007)
Log[Household Child Labor Use (Hours)/HA]	0.001	0.000	0.003	0.001	-0.003	-0.000
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)
Log[Hired Labor Use (Days)/HA]	0.079***	0.075***	0.074***	0.080***	0.079***	0.078***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Log[Exchange Labor Use (Days)/HA]	0.040***	0.026**	0.032***	0.038***	0.041***	0.039***
	(0.012)	(0.011)	(0.011)	(0.012)	(0.012)	(0.012)
Plot Location						
Elevation (M)	0.000***	0.000	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Table H.3 (Cont'd)

		Category of Additional Covariates Integrated into the Base Regression				
	Base	*District Fixed Effects*	*Plot Geospatial Characteristics*	*Other Plot Characteristics*	*Household Characteristics*	*Community Characteristics*
Plot Location						
Distance to Household (KM)	-0.001**	-0.001	-0.001	-0.001**	-0.001**	-0.001**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Plot Cultivation						
Intercropped †	0.110***	0.291***	0.179***	0.096***	0.111***	0.114***
	(0.025)	(0.027)	(0.025)	(0.026)	(0.025)	(0.025)
Share of Plot Area Under Improved Seeds	0.099***	0.079***	0.100***	0.094***	0.097***	0.100***
	(0.023)	(0.022)	(0.022)	(0.022)	(0.023)	(0.023)
Share of Plot Area Under Export Crops	1.213***	1.183***	1.218***	1.230***	1.213***	1.205***
	(0.040)	(0.040)	(0.040)	(0.041)	(0.039)	(0.040)
Household Characteristics						
Household Size	0.014***	0.011**	0.012***	0.013***	-0.067***	0.015***
	(0.005)	(0.004)	(0.004)	(0.005)	(0.024)	(0.005)
Child Dependency Ratio	-0.011	-0.014	-0.020	-0.008	-0.014	-0.012
	(0.016)	(0.015)	(0.015)	(0.016)	(0.022)	(0.016)
Agricultural Extension Receipt †	0.077***	0.034*	0.071***	0.075***	0.079***	0.082***
	(0.021)	(0.019)	(0.020)	(0.021)	(0.021)	(0.021)
Access to Non-Farm Labor Income †	-0.076***	-0.076***	-0.075***	-0.076***	-0.080***	-0.075***
	(0.019)	(0.017)	(0.018)	(0.019)	(0.019)	(0.019)
Access to Non-Farm Non-Labor Income †	-0.054**	-0.010	-0.027	-0.053*	-0.057**	-0.054**
	(0.027)	(0.025)	(0.024)	(0.028)	(0.027)	(0.027)
Wealth Index	0.055***	0.059***	0.055***	0.056***	0.055***	0.055***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Agricultural Implement Access Index	0.042***	0.030***	0.038***	0.041***	0.042***	0.041***
	(0.008)	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)
Distance to Nearest ADMARC (KM)	0.001	0.003	0.002	0.001	0.002	0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Household Agro-Ecological Zone Classification						
Tropic-warm/semiarid †	0.170**	0.017	0.190***	0.175**	0.171**	0.183***
	(0.071)	(0.088)	(0.069)	(0.072)	(0.074)	(0.065)
Tropic-warm/subhumid †	0.099	0.157*	0.158**	0.096	0.101	0.115
	(0.073)	(0.086)	(0.073)	(0.075)	(0.076)	(0.071)
Tropic-cool/semiarid †	0.073	-0.044	0.098	0.069	0.068	0.093
	(0.079)	(0.089)	(0.074)	(0.080)	(0.081)	(0.075)
Observations	16,372	16,372	16,016	16,153	16,372	16,234
Adjusted R-Squared	0.335	0.365	0.357	0.335	0.336	0.336

Table H.4: Exploring the Presence of Omitted Variable Bias in Base OLS Regression Results Underlying the Mean Decomposition

Dependent Variable: Log[Plot Value of Output (MK)/HA]

Male-Managed Plot Sample

		Category of Additional Covariates Integrated into the Base Regression				
	Base	*District Fixed Effects*	*Plot Geospatial Characteristics*	*Other Plot Characteristics*	*Household Characteristics*	*Community Characteristics*
Plot Manager Characteristics						
Female †						
Manager & Owner Overlap †	0.020	-0.005	0.024	0.018	0.021	0.018
	(0.022)	(0.022)	(0.022)	(0.023)	(0.022)	(0.022)
Age (Years)	-0.002*	-0.002*	-0.002*	-0.002**	0.000	-0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Years of Schooling	0.005	0.002	0.004	0.005	0.005	0.005
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Plot Area						
Log[GPS-Based Plot Area (HA)]	-0.261***	-0.267***	-0.260***	-0.262***	-0.262***	-0.268***
	(0.034)	(0.034)	(0.035)	(0.035)	(0.034)	(0.034)
Log[GPS-Based Plot Area (HA) Squared]	0.043***	0.046***	0.045***	0.044***	0.043***	0.042***
	(0.010)	(0.010)	(0.011)	(0.011)	(0.010)	(0.010)
Plot Non-Labor Input Use						
Incidence of Pesticide/Herbicide Use †	0.360***	0.405***	0.326***	0.358***	0.367***	0.363***
	(0.077)	(0.070)	(0.075)	(0.078)	(0.077)	(0.077)
Incidence of Organic Fertilizer Use †	0.054*	0.046	0.063**	0.056*	0.055*	0.052*
	(0.032)	(0.032)	(0.031)	(0.032)	(0.032)	(0.031)
Log[Inorganic Fertilizer Use (KG)/HA]	0.081***	0.084***	0.083***	0.080***	0.081***	0.081***
	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)
Plot Labor Input Use						
Log[Household Male Labor Use (Hours)/HA]	0.067***	0.066***	0.063***	0.067***	0.065***	0.069***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Log[Household Female Labor Use (Hours)/HA]	0.016**	0.015*	0.017**	0.017**	0.019**	0.017**
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Log[Household Child Labor Use (Hours)/HA]	0.006	0.007	0.007	0.006	0.004	0.005
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)
Log[Hired Labor Use (Days)/HA]	0.080***	0.076***	0.076***	0.081***	0.081***	0.080***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Log[Exchange Labor Use (Days)/HA]	0.042***	0.028**	0.036***	0.040***	0.044***	0.041***
	(0.014)	(0.014)	(0.013)	(0.014)	(0.014)	(0.014)
Plot Location						
Elevation (M)	0.000***	0.000	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Table H.4 (Cont'd)

		Category of Additional Covariates Integrated into the Base Regression				
	Base	*District Fixed Effects*	*Plot Geospatial Characteristics*	*Other Plot Characteristics*	*Household Characteristics*	*Community Characteristics*
Plot Location						
Distance to Household (KM)	-0.001	-0.000	0.000	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Plot Cultivation						
Intercropped †	0.089***	0.270***	0.156***	0.077***	0.089***	0.094***
	(0.028)	(0.032)	(0.029)	(0.029)	(0.028)	(0.029)
Share of Plot Area Under Improved Seeds	0.095***	0.078***	0.091***	0.088***	0.093***	0.098***
	(0.025)	(0.023)	(0.024)	(0.025)	(0.024)	(0.025)
Share of Plot Area Under Export Crops	1.187***	1.162***	1.191***	1.202***	1.183***	1.178***
	(0.040)	(0.040)	(0.040)	(0.042)	(0.040)	(0.040)
Household Characteristics						
Household Size	0.011*	0.008	0.008	0.010*	-0.061**	0.012**
	(0.006)	(0.006)	(0.006)	(0.006)	(0.028)	(0.006)
Child Dependency Ratio	0.032	0.024	0.024	0.033	0.039	0.029
	(0.020)	(0.019)	(0.020)	(0.021)	(0.028)	(0.021)
Agricultural Extension Receipt †	0.053**	0.015	0.048**	0.053**	0.054**	0.056**
	(0.022)	(0.021)	(0.022)	(0.022)	(0.022)	(0.023)
Access to Non-Farm Labor Income †	-0.075***	-0.076***	-0.081***	-0.073***	-0.076***	-0.071***
	(0.021)	(0.020)	(0.021)	(0.021)	(0.022)	(0.022)
Access to Non-Farm Non-Labor Income †	-0.035	0.004	-0.011	-0.028	-0.035	-0.031
	(0.031)	(0.029)	(0.029)	(0.032)	(0.031)	(0.030)
Wealth Index	0.062***	0.066***	0.061***	0.063***	0.062***	0.062***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Agricultural Implement Access Index	0.041***	0.029***	0.038***	0.041***	0.041***	0.040***
	(0.010)	(0.010)	(0.009)	(0.010)	(0.010)	(0.010)
Distance to Nearest ADMARC (KM)	0.004	0.005*	0.003	0.004	0.004	0.004
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Household Agro-Ecological Zone Classification						
Tropic-warm/semiarid †	0.186**	0.038	0.213***	0.196**	0.189**	0.215***
	(0.077)	(0.091)	(0.076)	(0.078)	(0.081)	(0.070)
Tropic-warm/subhumid †	0.135*	0.168*	0.174**	0.136*	0.136*	0.161**
	(0.081)	(0.091)	(0.080)	(0.082)	(0.082)	(0.077)
Tropic-cool/semiarid †	0.119	-0.005	0.158**	0.117	0.118	0.152*
	(0.086)	(0.093)	(0.080)	(0.087)	(0.088)	(0.082)
Observations	12,029	12,029	11,755	11,887	12,029	11,920
R-Squared	0.341	0.369	0.360	0.341	0.341	0.343

Table H.5: Exploring the Presence of Omitted Variable Bias in Base OLS Regression Results Underlying the Mean Decomposition

Dependent Variable: Log[Plot Value of Output (MK)/HA]

Female-Managed Plot Sample

		Category of Additional Covariates Integrated into the Base Regression				
	Base	*District Fixed Effects*	*Plot Geospatial Characteristics*	*Other Plot Characteristics*	*Household Characteristics*	*Community Characteristics*
Plot Manager Characteristics						
Female †						
Manager & Owner Overlap †	-0.015	0.026	-0.004	-0.009	-0.018	-0.014
	(0.040)	(0.038)	(0.038)	(0.041)	(0.040)	(0.040)
Age (Years)	0.001	-0.000	0.000	0.001	0.002	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Years of Schooling	0.015***	0.011**	0.015***	0.015***	0.015***	0.014***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Plot Area						
Log[GPS-Based Plot Area (HA)]	-0.296***	-0.326***	-0.322***	-0.292***	-0.288***	-0.299***
	(0.046)	(0.046)	(0.049)	(0.047)	(0.046)	(0.046)
Log[GPS-Based Plot Area (HA) Squared]	0.042***	0.039***	0.040***	0.043***	0.044***	0.041***
	(0.013)	(0.013)	(0.014)	(0.013)	(0.013)	(0.013)
Plot Non-Labor Input Use						
Incidence of Pesticide/Herbicide Use †	0.491***	0.526***	0.536***	0.471***	0.487***	0.501***
	(0.136)	(0.118)	(0.119)	(0.139)	(0.137)	(0.134)
Incidence of Organic Fertilizer Use †	0.017	0.030	0.020	0.022	0.013	0.008
	(0.045)	(0.044)	(0.043)	(0.046)	(0.046)	(0.046)
Log[Inorganic Fertilizer Use (KG)/HA]	0.066***	0.068***	0.070***	0.065***	0.065***	0.065***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Plot Labor Input Use						
Log[Household Male Labor Use (Hours)/HA]	0.005	0.004	0.003	0.006	0.007	0.006
	(0.007)	(0.006)	(0.006)	(0.007)	(0.008)	(0.007)
Log[Household Female Labor Use (Hours)/HA]	0.053***	0.050***	0.051***	0.056***	0.050***	0.054***
	(0.015)	(0.015)	(0.015)	(0.016)	(0.015)	(0.016)
Log[Household Child Labor Use (Hours)/HA]	-0.011*	-0.013**	-0.007	-0.011*	-0.015**	-0.011*
	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)
Log[Hired Labor Use (Days)/HA]	0.088***	0.087***	0.084***	0.090***	0.088***	0.090***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Log[Exchange Labor Use (Days)/HA]	0.033	0.023	0.018	0.030	0.029	0.032*
	(0.020)	(0.019)	(0.018)	(0.020)	(0.020)	(0.019)
Plot Location						
Elevation (M)	0.000***	0.000	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Table H.5 (Cont'd)

		Category of Additional Covariates Integrated into the Base Regression				
	Base	*District Fixed Effects*	*Plot Geospatial Characteristics*	*Other Plot Characteristics*	*Household Characteristics*	*Community Characteristics*
Plot Location						
Distance to Household (KM)	-0.002***	-0.002***	-0.003***	-0.002**	-0.002***	-0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Plot Cultivation						
Intercropped †	0.165***	0.344***	0.233***	0.146***	0.167***	0.169***
	(0.039)	(0.040)	(0.039)	(0.040)	(0.038)	(0.038)
Share of Plot Area Under Improved Seeds	0.099**	0.070*	0.107***	0.101**	0.098**	0.093**
	(0.041)	(0.040)	(0.037)	(0.041)	(0.042)	(0.043)
Share of Plot Area Under Export Crops	1.255***	1.212***	1.245***	1.273***	1.277***	1.257***
	(0.090)	(0.096)	(0.089)	(0.092)	(0.090)	(0.091)
Household Characteristics						
Household Size	0.033***	0.028***	0.028***	0.031***	-0.020	0.032***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.044)	(0.008)
Child Dependency Ratio	-0.076***	-0.080***	-0.090***	-0.072**	-0.047	-0.075**
	(0.029)	(0.027)	(0.026)	(0.029)	(0.032)	(0.029)
Agricultural Extension Receipt †	0.157***	0.103***	0.147***	0.149***	0.157***	0.168***
	(0.040)	(0.038)	(0.036)	(0.041)	(0.040)	(0.040)
Access to Non-Farm Labor Income †	-0.057	-0.048	-0.041	-0.059	-0.060	-0.062*
	(0.037)	(0.035)	(0.034)	(0.037)	(0.037)	(0.037)
Access to Non-Farm Non-Labor Income †	-0.097**	-0.045	-0.068*	-0.106**	-0.113***	-0.104**
	(0.043)	(0.038)	(0.040)	(0.044)	(0.042)	(0.043)
Wealth Index	0.048***	0.054***	0.051***	0.047***	0.043***	0.045***
	(0.011)	(0.011)	(0.011)	(0.012)	(0.011)	(0.011)
Agricultural Implement Access Index	0.043***	0.030**	0.042***	0.042***	0.042***	0.041***
	(0.014)	(0.014)	(0.012)	(0.014)	(0.014)	(0.014)
Distance to Nearest ADMARC (KM)	-0.005	-0.002	-0.000	-0.006	-0.004	-0.004
	(0.005)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
Household Agro-Ecological Zone Classification						
Tropic-warm/semiarid †	0.071	-0.127	0.096	0.053	0.071	0.026
	(0.085)	(0.138)	(0.085)	(0.090)	(0.088)	(0.087)
Tropic-warm/subhumid †	-0.035	0.046	0.078	-0.055	-0.022	-0.064
	(0.093)	(0.136)	(0.093)	(0.098)	(0.094)	(0.095)
Tropic-cool/semiarid †	-0.103	-0.217	-0.117	-0.115	-0.116	-0.127
	(0.090)	(0.135)	(0.090)	(0.096)	(0.095)	(0.092)
Observations	4,343	4,343	4,261	4,266	4,343	4,314
R-Squared	0.303	0.338	0.334	0.299	0.308	0.303

Table H.6: Within Household OLS Regression Results

Dependent Variable: Log[Plot Value of Output (MK)/HA]

	Base Pooled Sample	*Pooled Within Sample*	*Pooled Within Sample HHFE*
Plot Manager Characteristics			
Female †	-0.045*	-0.066	-0.114
	(0.027)	(0.097)	(0.113)
Manager & Owner Overlap †	0.016	0.155	0.306**
	(0.020)	(0.096)	(0.134)
Age (Years)	-0.001	0.005	0.002
	(0.001)	(0.004)	(0.004)
Years of Schooling	0.007**	0.029*	-0.017
	(0.003)	(0.016)	(0.023)
Plot Area			
Log[GPS-Based Plot Area (HA)]	-0.282***	-0.404**	-0.117
	(0.030)	(0.166)	(0.178)
Log[GPS-Based Plot Area (HA) Squared]	0.044***	0.008	0.061*
	(0.009)	(0.036)	(0.033)
Plot Non-Labor Input Use			
Incidence of Pesticide/Herbicide Use †	0.395***	1.837***	1.497***
	(0.076)	(0.280)	(0.350)
Incidence of Organic Fertilizer Use †	0.043	0.049	-0.017
	(0.027)	(0.178)	(0.205)
Log[Inorganic Fertilizer Use (KG)/HA]	0.077***	0.097***	0.076***
	(0.004)	(0.019)	(0.028)
Plot Labor Input Use			
Log[Household Male Labor Use (Hours)/HA]	0.028***	-0.025	0.016
	(0.005)	(0.028)	(0.033)
Log[Household Female Labor Use (Hours)/HA]	0.032***	0.050	0.048
	(0.007)	(0.036)	(0.032)
Log[Household Child Labor Use (Hours)/HA]	0.001	0.013	-0.001
	(0.004)	(0.023)	(0.029)
Log[Hired Labor Use (Days)/HA]	0.079***	0.048	0.075
	(0.008)	(0.034)	(0.061)
Log[Exchange Labor Use (Days)/HA]	0.040***	-0.045	-0.075
	(0.012)	(0.040)	(0.047)
Plot Location			
Elevation (M)	0.000***	0.000	-0.001
	(0.000)	(0.000)	(0.003)
Distance to Household (KM)	-0.001**	0.002	-0.022
	(0.001)	(0.007)	(0.032)

Table H.6 (Cont'd)

	Base Pooled Sample	*Pooled Within Sample*	*Pooled Within Sample HHFE*
Plot Cultivation			
Intercropped †	0.110***	0.126	0.233
	(0.025)	(0.142)	(0.155)
Share of Plot Area Under Improved Seeds	0.099***	0.247*	0.235
	(0.023)	(0.147)	(0.173)
Share of Plot Area Under Export Crops	1.213***	1.316***	1.197***
	(0.040)	(0.168)	(0.131)
Household Characteristics			
Household Size	0.014***	0.031	
	(0.005)	(0.026)	
Child Dependency Ratio	-0.011	-0.007	
	(0.016)	(0.117)	
Agricultural Extension Receipt †	0.077***	-0.152	
	(0.021)	(0.157)	
Household Characteristics			
Access to Non-Farm Labor Income †	-0.076***	-0.120	
	(0.019)	(0.112)	
Access to Non-Farm Non-Labor Income †	-0.054**	-0.229	
	(0.027)	(0.142)	
Wealth Index	0.055***	-0.023	
	(0.006)	(0.028)	
Agricultural Implement Access Index	0.042***	0.111*	
	(0.008)	(0.061)	
Distance to Nearest ADMARC (KM)	0.001	-0.013	
	(0.003)	(0.018)	
Household Agro-Ecological Zone Classification			
Tropic-warm/semiarid †	0.170**	-0.013	
	(0.071)	(0.231)	
Tropic-warm/subhumid †	0.099	-0.017	
	(0.073)	(0.276)	
Tropic-cool/semiarid †	0.073	-0.312	
	(0.079)	(0.249)	
Observations	16,372	292	292
R-Squared	0.336	0.430	0.707

Note: The estimates are weighted in accordance with the complex survey design.
***/**/* indicate statistical significance at the 1/5/10 percent level, respectively.
†denotes a dummy variable.

Table H.7: Decomposition of the Gender Differential in Agricultural Productivity - Within Household Sample
Agricultural Productivity Proxied by Log[Plot Gross Value of Output (MK)/HA]

A. Mean Gender Differential			
Mean Male-Managed Plot Agricultural Productivity	10.666***		
	(0.085)		
Mean Female-Managed Plot Agricultural Productivity	10.392***		
	(0.086)		
Mean Gender Differential in Agricultural Productivty	0.274**		
	(0.112)		

B. Aggregate Decomposition	*Endowment Effect*	*Male Structural Advantage*	*Female Structural Disadvantage*
TOTAL	0.208**	-0.000	0.066
	(0.093)	(0.027)	(0.093)

C. Detailed Decomposition	*Endowment Effect*	*Male Structural Advantage*	*Female Structural Disadvantage*
Plot Manager Characteristics			
Manager & Owner Overlap †	0.013	-0.009	0.001
	(0.013)	(0.072)	(0.067)
Age (Years)	0.026	-0.069	-0.109
	(0.023)	(0.140)	(0.204)
Years of Schooling	0.034	-0.049	-0.148*
	(0.021)	(0.071)	(0.085)
Plot Area			
Log[GPS-Based Plot Area (HA)]	-0.096	-0.348	-0.292
	(0.059)	(0.225)	(0.327)
Log[GPS-Based Plot Area (HA) Squared]	-0.007	0.236*	0.205
	(0.031)	(0.126)	(0.152)
Plot Non-Labor Input Use			
Incidence of Pesticide/Herbicide Use †	-0.008	0.001	0.004
	(0.019)	(0.003)	(0.008)
Incidence of Organic Fertilizer Use †	0.003	-0.008	0.021
	(0.012)	(0.023)	(0.018)
Log[Inorganic Fertilizer Use (KG)/HA]	0.099***	0.066	0.042
	(0.037)	(0.066)	(0.080)
Plot Labor Input Use			
Log[Household Male Labor Use (Hours)/HA]	-0.006	0.104	0.200
	(0.008)	(0.155)	(0.172)
Log[Household Female Labor Use (Hours)/HA]	-0.008	0.155	-0.063
	(0.011)	(0.217)	(0.264)
Log[Household Child Labor Use (Hours)/HA]	0.000	-0.011	-0.029
	(0.002)	(0.028)	(0.033)

Table H.7 (Cont'd)			
C. Detailed Decomposition (Cont'd)	*Endowment Effect*	*Male Structural Advantage*	*Female Structural Disadvantage*
Log[Hired Labor Use (Days)/HA]	-0.005	0.075**	0.108**
	(0.008)	(0.036)	(0.053)
Log[Exchange Labor Use (Days)/HA]	-0.005	0.009	-0.006
	(0.008)	(0.017)	(0.017)
Plot Location			
Elevation (M)	0.008	-0.052	-0.155
	(0.021)	(0.221)	(0.278)
Distance to Household (KM)	-0.001	0.004	0.004
	(0.004)	(0.015)	(0.017)
Plot Cultivation			
Intercropped †	-0.014	0.012	-0.014
	(0.016)	(0.031)	(0.046)
Share of Plot Area Under Improved Seeds	-0.011	-0.024	-0.053
	(0.019)	(0.036)	(0.054)
Share of Plot Area Under Export Crops	0.158***	0.002	-0.003
	(0.054)	(0.014)	(0.007)
Household Characteristics			
Household Size	-0.004	-0.195	-0.288
	(0.005)	(0.132)	(0.202)
Child Dependency Ratio	-0.000	0.069	0.072
	(0.000)	(0.075)	(0.097)
Agricultural Extension Receipt †	-0.003	-0.007	0.015
	(0.005)	(0.021)	(0.032)
Access to Non-Farm Labor Income †	0.008	-0.046	-0.078
	(0.008)	(0.039)	(0.072)
Access to Non-Farm Non-Labor Income †	0.013	0.006	-0.019
	(0.010)	(0.038)	(0.062)
Wealth Index	-0.002	0.005	0.007
	(0.003)	(0.011)	(0.019)
Agricultural Implement Access Index	0.015	-0.049	-0.053
	(0.010)	(0.037)	(0.044)
Distance to Nearest ADMARC (KM)	-0.005	-0.015	0.003
	(0.007)	(0.081)	(0.091)
Household Agro-Ecological Zone Classification [Aggregated]	0.007	0.024	-0.021
	(0.016)	(0.028)	(0.044)
Observations		292	

Note: The estimates are weighted in accordance with the complex survey design.
***/**/* indicate statistical significance at the 1/5/10 percent level, respectively.
†denotes a dummy variable.

Table H.8: Base OLS Regression Results Underlying the RIF Decomposition at the 10th Percentile

Dependent Variable: RIF(Log[Plot Value of Output (MK)/HA])

	Pooled Sample	*Male-Managed Plot Sample*	*Female-Managed Plot Sample*
Plot Manager Characteristics			
Female †	0.008		
	(0.052)		
Manager & Owner Overlap †	0.086**	0.093**	0.085
	(0.036)	(0.042)	(0.084)
Age (Years)	-0.000	-0.002	0.001
	(0.001)	(0.002)	(0.003)
Years of Schooling	0.009*	0.005	0.028***
	(0.005)	(0.006)	(0.010)
Plot Area			
Log[GPS-Based Plot Area (HA)]	-0.617***	-0.572***	-0.741***
	(0.060)	(0.072)	(0.113)
Log[GPS-Based Plot Area (HA) Squared]	-0.064***	-0.063***	-0.078***
	(0.015)	(0.019)	(0.024)
Plot Non-Labor Input Use			
Incidence of Pesticide/Herbicide Use †	0.331***	0.269**	0.601**
	(0.121)	(0.135)	(0.292)
Incidence of Organic Fertilizer Use †	0.081	0.072	0.124
	(0.053)	(0.063)	(0.102)
Log[Inorganic Fertilizer Use (KG)/HA]	0.132***	0.133***	0.141***
	(0.008)	(0.009)	(0.016)
Plot Labor Input Use			
Log[Household Male Labor Use (Hours)/HA]	0.051***	0.076***	0.043***
	(0.009)	(0.017)	(0.012)
Log[Household Female Labor Use (Hours)/HA]	0.052***	0.029**	0.112***
	(0.012)	(0.013)	(0.031)
Log[Household Child Labor Use (Hours)/HA]	-0.004	0.006	-0.009
	(0.007)	(0.009)	(0.013)
Log[Hired Labor Use (Days)/HA]	0.065***	0.061***	0.084***
	(0.012)	(0.015)	(0.024)
Log[Exchange Labor Use (Days)/HA]	0.056***	0.051**	0.038
	(0.020)	(0.025)	(0.035)
Plot Location			
Elevation (M)	0.000***	0.000***	0.001***
	(0.000)	(0.000)	(0.000)
Distance to Household (KM)	-0.001	-0.001	0.000
	(0.002)	(0.002)	(0.002)
Plot Cultivation			
Intercropped †	0.076*	0.022	0.197**
	(0.041)	(0.050)	(0.077)
Share of Plot Area Under Improved Seeds	0.121***	0.169***	-0.010
	(0.043)	(0.050)	(0.090)
Share of Plot Area Under Export Crops	0.247***	0.306***	0.055
	(0.043)	(0.048)	(0.145)

Table H.8 (Cont'd)

	Pooled Sample	*Male-Managed Plot Sample*	*Female-Managed Plot Sample*
Household Characteristics			
Household Size	0.021**	0.017	0.042**
	(0.009)	(0.011)	(0.017)
Child Dependency Ratio	-0.003	0.085**	-0.067
	(0.034)	(0.038)	(0.052)
Agricultural Extension Receipt †	0.125***	0.081*	0.163**
	(0.035)	(0.041)	(0.072)
Household Characteristics			
Access to Non-Farm Labor Income †	-0.098***	-0.048	-0.267***
	(0.035)	(0.041)	(0.075)
Access to Non-Farm Non-Labor Income †	-0.141***	-0.108**	-0.208**
	(0.043)	(0.051)	(0.088)
Wealth Index	0.050***	0.056***	0.044***
	(0.009)	(0.010)	(0.015)
Agricultural Implement Access Index	0.042***	0.033**	0.086***
	(0.014)	(0.016)	(0.029)
Distance to Nearest ADMARC (KM)	-0.002	0.004	-0.014**
	(0.003)	(0.004)	(0.007)
Household Agro-Ecological Zone Classification			
Tropic-warm/semiarid †	0.416***	0.436***	0.443***
	(0.079)	(0.091)	(0.163)
Tropic-warm/subhumid †	0.168**	0.273***	0.021
	(0.083)	(0.096)	(0.170)
Tropic-cool/semiarid †	0.111	0.155	-0.029
	(0.085)	(0.098)	(0.181)
Observations	16,372	12,029	4,343
R-Squared	0.092	0.089	0.109

Note: The estimates are weighted in accordance with the complex survey design. ***/**/* indicate statistical significance at the 1/5/10 percent level, respectively. †denotes a dummy variable.

Table H.9: Base OLS Regression Results Underlying the RIF Decomposition at the 50th Percentile

Dependent Variable: RIF(Log[Plot Value of Output (MK)/HA])

	Pooled Sample	*Male-Managed Plot Sample*	*Female-Managed Plot Sample*
Plot Manager Characteristics			
Female †	-0.075**		
	(0.030)		
Manager & Owner Overlap †	0.019	0.012	-0.039
	(0.020)	(0.023)	(0.041)
Age (Years)	-0.000	-0.001	0.002*
	(0.001)	(0.001)	(0.001)
Years of Schooling	0.009***	0.009***	0.014**
	(0.003)	(0.003)	(0.006)
Plot Area			
Log[GPS-Based Plot Area (HA)]	-0.277***	-0.247***	-0.349***
	(0.025)	(0.030)	(0.049)
Log[GPS-Based Plot Area (HA) Squared]	0.032***	0.039***	0.011
	(0.007)	(0.009)	(0.011)
Plot Non-Labor Input Use			
Incidence of Pesticide/Herbicide Use †	0.348***	0.305***	0.560***
	(0.064)	(0.069)	(0.139)
Incidence of Organic Fertilizer Use †	0.051*	0.059*	0.050
	(0.029)	(0.033)	(0.052)
Log[Inorganic Fertilizer Use (KG)/HA]	0.075***	0.077***	0.066***
	(0.004)	(0.005)	(0.007)
Plot Labor Input Use			
Log[Household Male Labor Use (Hours)/HA]	0.025***	0.067***	0.010
	(0.005)	(0.010)	(0.007)
Log[Household Female Labor Use (Hours)/HA]	0.030***	0.010	0.041**
	(0.007)	(0.008)	(0.017)
Log[Household Child Labor Use (Hours)/HA]	-0.002	0.006	-0.013*
	(0.004)	(0.005)	(0.007)
Log[Hired Labor Use (Days)/HA]	0.076***	0.075***	0.075***
	(0.008)	(0.010)	(0.015)
Log[Exchange Labor Use (Days)/HA]	0.047***	0.045***	0.031
	(0.013)	(0.017)	(0.020)
Plot Location			
Elevation (M)	0.000***	0.000***	0.000**
	(0.000)	(0.000)	(0.000)
Distance to Household (KM)	-0.002***	-0.002	-0.003***
	(0.001)	(0.001)	(0.001)
Plot Cultivation			
Intercropped †	0.106***	0.121***	0.160***
	(0.022)	(0.027)	(0.038)
Share of Plot Area Under Improved Seeds	0.121***	0.123***	0.133***
	(0.024)	(0.028)	(0.043)
Share of Plot Area Under Export Crops	0.905***	0.944***	0.864***
	(0.029)	(0.033)	(0.070)

Table H.9 (Cont'd)

	Pooled Sample	*Male-Managed Plot Sample*	*Female-Managed Plot Sample*
Household Characteristics			
Household Size	0.007	0.008	0.025***
	(0.005)	(0.006)	(0.009)
Child Dependency Ratio	0.005	0.031	-0.040
	(0.017)	(0.023)	(0.025)
Agricultural Extension Receipt †	0.068***	0.058**	0.188***
	(0.021)	(0.024)	(0.039)
Household Characteristics			
Access to Non-Farm Labor Income †	-0.060***	-0.063***	-0.022
	(0.020)	(0.023)	(0.036)
Access to Non-Farm Non-Labor Income †	-0.044*	-0.007	-0.137***
	(0.024)	(0.028)	(0.042)
Wealth Index	0.061***	0.070***	0.045***
	(0.006)	(0.007)	(0.010)
Agricultural Implement Access Index	0.045***	0.041***	0.046***
	(0.008)	(0.009)	(0.014)
Distance to Nearest ADMARC (KM)	0.004**	0.004**	-0.002
	(0.002)	(0.002)	(0.003)
Household Agro-Ecological Zone Classification			
Tropic-warm/semiarid †	0.181***	0.185***	0.036
	(0.045)	(0.051)	(0.094)
Tropic-warm/subhumid †	0.117**	0.108**	-0.029
	(0.047)	(0.053)	(0.096)
Tropic-cool/semiarid †	0.090**	0.145***	-0.102
	(0.046)	(0.052)	(0.096)
Observations	16,372	12,029	4,343
R-Squared	0.205	0.213	0.193

Note: The estimates are weighted in accordance with the complex survey design. ***/**/* indicate statistical significance at the 1/5/10 percent level, respectively. †denotes a dummy variable.

Table H.10: Base OLS Regression Results Underlying the RIF Decomposition at the 90th Percentile

Dependent Variable: RIF(Log[Plot Value of Output (MK)/HA])

	Pooled Sample	*Male-Managed Plot Sample*	*Female-Managed Plot Sample*
Plot Manager Characteristics			
Female †	-0.082**		
	(0.042)		
Manager & Owner Overlap †	-0.026	-0.021	-0.088
	(0.033)	(0.036)	(0.083)
Age (Years)	-0.002	-0.001	-0.002
	(0.001)	(0.001)	(0.002)
Years of Schooling	0.001	-0.000	0.019*
	(0.005)	(0.005)	(0.011)
Plot Area			
Log[GPS-Based Plot Area (HA)]	0.065	0.101*	0.035
	(0.054)	(0.059)	(0.105)
Log[GPS-Based Plot Area (HA) Squared]	0.188***	0.183***	0.213***
	(0.021)	(0.024)	(0.033)
Plot Non-Labor Input Use			
Incidence of Pesticide/Herbicide Use †	0.703***	0.765***	0.232
	(0.137)	(0.154)	(0.320)
Incidence of Organic Fertilizer Use †	0.077	0.127*	-0.038
	(0.055)	(0.065)	(0.094)
Log[Inorganic Fertilizer Use (KG)/HA]	0.035***	0.036***	0.023
	(0.006)	(0.007)	(0.014)
Plot Labor Input Use			
Log[Household Male Labor Use (Hours)/HA]	0.020**	0.084***	-0.034***
	(0.009)	(0.017)	(0.013)
Log[Household Female Labor Use (Hours)/HA]	0.031**	0.005	0.067**
	(0.014)	(0.016)	(0.032)
Log[Household Child Labor Use (Hours)/HA]	0.000	0.002	-0.001
	(0.007)	(0.008)	(0.013)
Log[Hired Labor Use (Days)/HA]	0.104***	0.117***	0.144***
	(0.015)	(0.017)	(0.033)
Log[Exchange Labor Use (Days)/HA]	0.013	0.015	0.022
	(0.020)	(0.025)	(0.039)
Plot Location			
Elevation (M)	-0.000*	-0.000**	0.000
	(0.000)	(0.000)	(0.000)
Distance to Household (KM)	0.001	0.002	-0.001
	(0.001)	(0.002)	(0.001)
Plot Cultivation			
Intercropped †	0.135***	0.097**	0.152**
	(0.034)	(0.040)	(0.070)
Share of Plot Area Under Improved Seeds	0.023	-0.010	-0.022
	(0.031)	(0.034)	(0.073)
Share of Plot Area Under Export Crops	3.062***	2.950***	3.521***
	(0.107)	(0.114)	(0.336)

Table H.10 (Cont'd)

	Pooled Sample	*Male-Managed Plot Sample*	*Female-Managed Plot Sample*
Household Characteristics			
Household Size	0.006	0.005	0.054***
	(0.008)	(0.010)	(0.017)
Child Dependency Ratio	-0.006	0.029	-0.121***
	(0.026)	(0.037)	(0.042)
Agricultural Extension Receipt †	0.052	0.039	0.147*
	(0.034)	(0.038)	(0.076)
Household Characteristics			
Access to Non-Farm Labor Income †	-0.130***	-0.116***	-0.049
	(0.031)	(0.036)	(0.067)
Access to Non-Farm Non-Labor Income †	-0.016	0.004	-0.043
	(0.039)	(0.045)	(0.084)
Wealth Index	0.060***	0.056***	0.069***
	(0.010)	(0.011)	(0.023)
Agricultural Implement Access Index	0.036***	0.033**	0.045*
	(0.013)	(0.014)	(0.025)
Distance to Nearest ADMARC (KM)	0.003	0.007**	-0.011**
	(0.003)	(0.003)	(0.005)
Household Agro-Ecological Zone Classification			
Tropic-warm/semiarid †	-0.200***	-0.142*	-0.185
	(0.072)	(0.075)	(0.188)
Tropic-warm/subhumid †	-0.146*	-0.062	-0.178
	(0.076)	(0.079)	(0.194)
Tropic-cool/semiarid †	-0.114	-0.011	-0.298
	(0.073)	(0.076)	(0.190)
Observations	16,372	12,029	4,343
R-Squared	0.260	0.273	0.210

Note: The estimates are weighted in accordance with the complex survey design.
***/**/* indicate statistical significance at the 1/5/10 percent level, respectively.
†denotes a dummy variable.

Table H.11: Exploring the Presence of Omitted Variable Bias in Base OLS Regression Results Underlying the RIF Decomposition at the 10th

Dependent Variable: RIF(Log[Plot Value of Output (MK)/HA])

Pooled Sample

		Category of Additional Covariates Integrated into the Base Regression				
	Base	*District Fixed Effects*	*Plot Geospatial Characteristics*	*Other Plot Characteristics*	*Household Characteristics*	*Community Characteristics*
Plot Manager Characteristics						
Female †	0.008	-0.012	-0.009	0.001	0.006	0.023
	(0.052)	(0.051)	(0.052)	(0.052)	(0.052)	(0.051)
Manager & Owner Overlap †	0.086**	0.056	0.086**	0.076**	0.085**	0.080**
	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)
Age (Years)	-0.000	-0.001	-0.000	-0.001	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Years of Schooling	0.009*	0.005	0.007	0.009*	0.009*	0.008
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Plot Area						
Log[GPS-Based Plot Area (HA)]	-0.617***	-0.637***	-0.622***	-0.616***	-0.613***	-0.636***
	(0.060)	(0.060)	(0.061)	(0.061)	(0.060)	(0.060)
Log[GPS-Based Plot Area (HA) Squared]	-0.064***	-0.065***	-0.063***	-0.063***	-0.064***	-0.068***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Plot Non-Labor Input Use						
Incidence of Pesticide/Herbicide Use †	0.331***	0.484***	0.325***	0.299**	0.329***	0.327***
	(0.121)	(0.119)	(0.119)	(0.123)	(0.121)	(0.120)
Incidence of Organic Fertilizer Use †	0.081	0.087*	0.096*	0.108**	0.082	0.093*
	(0.053)	(0.053)	(0.053)	(0.049)	(0.053)	(0.053)
Log[Inorganic Fertilizer Use (KG)/HA]	0.132***	0.134***	0.134***	0.133***	0.132***	0.131***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Plot Labor Input Use						
Log[Household Male Labor Use (Hours)/HA]	0.051***	0.048***	0.047***	0.052***	0.055***	0.053***
	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)	(0.009)
Log[Household Female Labor Use (Hours)/HA]	0.052***	0.051***	0.050***	0.052***	0.052***	0.048***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.011)
Log[Household Child Labor Use (Hours)/HA]	-0.004	-0.005	-0.000	-0.003	-0.009	-0.005
	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)	(0.007)
Log[Hired Labor Use (Days)/HA]	0.065***	0.060***	0.060***	0.064***	0.066***	0.065***
	(0.012)	(0.012)	(0.012)	(0.013)	(0.012)	(0.012)
Log[Exchange Labor Use (Days)/HA]	0.056***	0.037*	0.047**	0.054***	0.056***	0.058***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
Plot Location						
Elevation (M)	0.000***	0.000	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Distance to Household (KM)	-0.001	-0.001	-0.001	-0.002	-0.001	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)

Table H.11 (Cont'd)

		Category of Additional Covariates Integrated into the Base Regression				
	Base	*District Fixed Effects*	*Plot Geospatial Characteristics*	*Other Plot Characteristics*	*Household Characteristics*	*Community Characteristics*
Plot Cultivation						
Intercropped †	0.076*	0.368***	0.173***	0.044	0.079*	0.083**
	(0.041)	(0.051)	(0.044)	(0.042)	(0.041)	(0.042)
Distance to Household (KM)	-0.001	-0.001	-0.001	-0.002	-0.001	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Plot Cultivation						
Intercropped †	0.076*	0.368***	0.173***	0.044	0.079*	0.083**
	(0.041)	(0.051)	(0.044)	(0.042)	(0.041)	(0.042)
Share of Plot Area Under Improved Seeds	0.121***	0.082*	0.117***	0.117***	0.117***	0.123***
	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)
Share of Plot Area Under Export Crops	0.247***	0.219***	0.247***	0.312***	0.245***	0.242***
	(0.043)	(0.043)	(0.044)	(0.047)	(0.043)	(0.043)
Household Characteristics						
Household Size	0.021**	0.013	0.015*	0.019**	-0.004	0.021**
	(0.009)	(0.009)	(0.009)	(0.009)	(0.045)	(0.009)
Child Dependency Ratio	-0.003	-0.013	-0.012	-0.005	-0.011	0.001
	(0.034)	(0.034)	(0.034)	(0.035)	(0.048)	(0.034)
Agricultural Extension Receipt †	0.125***	0.046	0.107***	0.116***	0.123***	0.125***
	(0.035)	(0.036)	(0.036)	(0.035)	(0.035)	(0.035)
Access to Non-Farm Labor Income †	-0.098***	-0.090**	-0.092***	-0.094***	-0.099***	-0.101***
	(0.035)	(0.035)	(0.036)	(0.035)	(0.035)	(0.035)
Access to Non-Farm Non-Labor Income †	-0.141***	-0.065	-0.091**	-0.139***	-0.147***	-0.132***
	(0.043)	(0.044)	(0.044)	(0.042)	(0.043)	(0.043)
Wealth Index	0.050***	0.054***	0.050***	0.051***	0.050***	0.048***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Agricultural Implement Access Index	0.042***	0.033**	0.033**	0.041***	0.042***	0.043***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Distance to Nearest ADMARC (KM)	-0.002	-0.001	-0.003	-0.003	-0.002	-0.001
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
Household Agro-Ecological Zone Classification						
Tropic-warm/semiarid †	0.416***	0.141	0.441***	0.434***	0.432***	0.421***
	(0.079)	(0.108)	(0.087)	(0.080)	(0.080)	(0.081)
Tropic-warm/subhumid †	0.168**	0.194*	0.274***	0.167**	0.181**	0.180**
	(0.083)	(0.109)	(0.087)	(0.084)	(0.084)	(0.086)
Tropic-cool/semiarid †	0.111	0.059	0.157*	0.121	0.120	0.143
	(0.085)	(0.113)	(0.092)	(0.086)	(0.086)	(0.087)
Observations	16,372	16,372	16,016	16,153	16,372	16,234
Adjusted R-Squared	0.091	0.113	0.099	0.091	0.091	0.091

Table H.12: Exploring the Presence of Omitted Variable Bias in Base OLS Regression Results Underlying the RIF Decomposition at the 10th

Dependent Variable: RIF(Log[Plot Value of Output (MK)/HA])

Male-Managed Plot Sample

		Category of Additional Covariates Integrated into the Base Regression				
	Base	*District Fixed Effects*	*Plot Geospatial Characteristics*	*Other Plot Characteristics*	*Household Characteristics*	*Community Characteristics*
Plot Manager Characteristics						
Female †						
Manager & Owner Overlap †	0.093**	0.058	0.088**	0.083**	0.092**	0.086**
	(0.042)	(0.042)	(0.043)	(0.042)	(0.042)	(0.042)
Age (Years)	-0.002	-0.002	-0.002	-0.002	-0.003	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Years of Schooling	0.005	0.003	0.005	0.005	0.005	0.006
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Plot Area						
Log[GPS-Based Plot Area (HA)]	-0.572***	-0.569***	-0.561***	-0.581***	-0.572***	-0.591***
	(0.072)	(0.074)	(0.074)	(0.074)	(0.072)	(0.073)
Log[GPS-Based Plot Area (HA) Squared]	-0.063***	-0.060***	-0.058***	-0.063***	-0.064***	-0.066***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
Plot Non-Labor Input Use						
Incidence of Pesticide/Herbicide Use †	0.269**	0.414***	0.256*	0.228	0.264*	0.287**
	(0.135)	(0.134)	(0.133)	(0.140)	(0.136)	(0.135)
Incidence of Organic Fertilizer Use †	0.072	0.066	0.080	0.101*	0.073	0.088
	(0.063)	(0.062)	(0.063)	(0.059)	(0.063)	(0.063)
Log[Inorganic Fertilizer Use (KG)/HA]	0.133***	0.134***	0.138***	0.133***	0.133***	0.135***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Plot Labor Input Use						
Log[Household Male Labor Use (Hours)/HA]	0.076***	0.076***	0.073***	0.078***	0.080***	0.079***
	(0.017)	(0.017)	(0.017)	(0.018)	(0.018)	(0.017)
Log[Household Female Labor Use (Hours)/HA]	0.029**	0.028**	0.028**	0.029**	0.027**	0.026*
	(0.013)	(0.013)	(0.013)	(0.014)	(0.014)	(0.013)
Log[Household Child Labor Use (Hours)/HA]	0.006	0.006	0.008	0.008	0.003	0.005
	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)	(0.009)
Log[Hired Labor Use (Days)/HA]	0.061***	0.056***	0.055***	0.061***	0.061***	0.060***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Log[Exchange Labor Use (Days)/HA]	0.051**	0.031	0.036	0.050*	0.051**	0.050*
	(0.025)	(0.026)	(0.026)	(0.026)	(0.025)	(0.025)
Plot Location						
Elevation (M)	0.000***	0.000	0.001***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Distance to Household (KM)	-0.001	-0.001	-0.000	-0.002	-0.001	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)

Table H.12 (Cont'd)

		Category of Additional Covariates Integrated into the Base Regression				
	Base	*District Fixed Effects*	*Plot Geospatial Characteristics*	*Other Plot Characteristics*	*Household Characteristics*	*Community Characteristics*
Plot Cultivation						
Intercropped †	0.022	0.288***	0.108**	0.002	0.024	0.026
	(0.050)	(0.062)	(0.053)	(0.051)	(0.050)	(0.050)
Distance to Household (KM)	-0.001	-0.001	-0.000	-0.002	-0.001	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Plot Cultivation						
Intercropped †	0.022	0.288***	0.108**	0.002	0.024	0.026
	(0.050)	(0.062)	(0.053)	(0.051)	(0.050)	(0.050)
Share of Plot Area Under Improved Seeds	0.169***	0.137***	0.149***	0.166***	0.168***	0.162***
	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)
Share of Plot Area Under Export Crops	0.306***	0.290***	0.292***	0.373***	0.303***	0.286***
	(0.048)	(0.048)	(0.048)	(0.053)	(0.047)	(0.048)
Household Characteristics						
Household Size	0.017	0.011	0.011	0.016	0.022	0.017
	(0.011)	(0.011)	(0.011)	(0.011)	(0.057)	(0.011)
Child Dependency Ratio	0.085**	0.073*	0.083**	0.081**	0.104*	0.083**
	(0.038)	(0.038)	(0.038)	(0.038)	(0.054)	(0.038)
Agricultural Extension Receipt †	0.081*	0.019	0.063	0.079*	0.077*	0.085**
	(0.041)	(0.042)	(0.042)	(0.042)	(0.041)	(0.041)
Access to Non-Farm Labor Income †	-0.048	-0.047	-0.046	-0.048	-0.048	-0.042
	(0.041)	(0.040)	(0.041)	(0.041)	(0.041)	(0.041)
Access to Non-Farm Non-Labor Income †	-0.108**	-0.036	-0.068	-0.099**	-0.115**	-0.098*
	(0.051)	(0.052)	(0.052)	(0.050)	(0.051)	(0.050)
Wealth Index	0.056***	0.058***	0.055***	0.059***	0.055***	0.052***
	(0.010)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Agricultural Implement Access Index	0.033**	0.024	0.031*	0.031*	0.034**	0.037**
	(0.016)	(0.017)	(0.016)	(0.017)	(0.016)	(0.017)
Distance to Nearest ADMARC (KM)	0.004	0.005	0.002	0.004	0.004	0.005
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Household Agro-Ecological Zone Classification						
Tropic-warm/semiarid †	0.436***	0.187	0.495***	0.476***	0.450***	0.456***
	(0.091)	(0.126)	(0.101)	(0.094)	(0.092)	(0.094)
Tropic-warm/subhumid †	0.273***	0.269**	0.353***	0.292***	0.283***	0.296***
	(0.096)	(0.126)	(0.101)	(0.100)	(0.097)	(0.099)
Tropic-cool/semiarid †	0.155	0.064	0.221**	0.177*	0.167*	0.194*
	(0.098)	(0.128)	(0.105)	(0.101)	(0.099)	(0.100)
Observations	12,029	12,029	11,755	11,887	12,029	11,920
Adjusted R-Squared	0.087	0.104	0.096	0.086	0.087	0.088

Table H.13: Exploring the Presence of Omitted Variable Bias in Base OLS Regression Results Underlying the RIF Decomposition at the 10th

Dependent Variable: RIF(Log[Plot Value of Output (MK)/HA])

Female-Managed Plot Sample

		Category of Additional Covariates Integrated into the Base Regression				
	Base	*District Fixed Effects*	*Plot Geospatial Characteristics*	*Other Plot Characteristics*	*Household Characteristics*	*Community Characteristics*
Plot Manager Characteristics						
Female †						
Manager & Owner Overlap †	0.085	0.130	0.124	0.073	0.085	0.101
	(0.084)	(0.086)	(0.086)	(0.083)	(0.084)	(0.086)
Age (Years)	0.001	-0.001	-0.000	0.001	0.001	0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Years of Schooling	0.028***	0.019*	0.023**	0.024**	0.027***	0.025**
	(0.010)	(0.010)	(0.011)	(0.010)	(0.010)	(0.011)
Plot Area						
Log[GPS-Based Plot Area (HA)]	-0.741***	-0.815***	-0.815***	-0.749***	-0.735***	-0.785***
	(0.113)	(0.109)	(0.114)	(0.110)	(0.112)	(0.114)
Log[GPS-Based Plot Area (HA) Squared]	-0.078***	-0.088***	-0.087***	-0.080***	-0.075***	-0.084***
	(0.024)	(0.023)	(0.025)	(0.024)	(0.024)	(0.024)
Plot Non-Labor Input Use						
Incidence of Pesticide/Herbicide Use †	0.601**	0.750***	0.659**	0.535*	0.600**	0.609**
	(0.292)	(0.288)	(0.281)	(0.290)	(0.292)	(0.298)
Incidence of Organic Fertilizer Use †	0.124	0.139	0.147	0.133	0.119	0.116
	(0.102)	(0.106)	(0.102)	(0.100)	(0.103)	(0.104)
Log[Inorganic Fertilizer Use (KG)/HA]	0.141***	0.143***	0.143***	0.140***	0.139***	0.140***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Plot Labor Input Use						
Log[Household Male Labor Use (Hours)/HA]	0.043***	0.040***	0.036***	0.045***	0.047***	0.046***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.014)	(0.012)
Log[Household Female Labor Use (Hours)/HA]	0.112***	0.106***	0.105***	0.090***	0.103***	0.095***
	(0.031)	(0.030)	(0.030)	(0.027)	(0.031)	(0.028)
Log[Household Child Labor Use (Hours)/HA]	-0.009	-0.012	-0.003	-0.011	-0.013	-0.010
	(0.013)	(0.013)	(0.013)	(0.013)	(0.015)	(0.014)
Log[Hired Labor Use (Days)/HA]	0.084***	0.079***	0.078***	0.080***	0.084***	0.080***
	(0.024)	(0.025)	(0.024)	(0.024)	(0.024)	(0.025)
Log[Exchange Labor Use (Days)/HA]	0.038	0.024	0.020	0.033	0.032	0.036
	(0.035)	(0.035)	(0.035)	(0.034)	(0.035)	(0.035)
Plot Location						
Elevation (M)	0.001***	-0.000	0.000**	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Distance to Household (KM)	0.000	0.000	0.000	0.001	0.000	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)

Table H.13 (Cont'd)

		Category of Additional Covariates Integrated into the Base Regression				
	Base	*District Fixed Effects*	*Plot Geospatial Characteristics*	*Other Plot Characteristics*	*Household Characteristics*	*Community Characteristics*
Plot Cultivation						
Intercropped †	0.197**	0.578***	0.326***	0.140*	0.198***	0.222***
	(0.077)	(0.097)	(0.084)	(0.078)	(0.077)	(0.080)
Distance to Household (KM)	0.000	0.000	0.000	0.001	0.000	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Plot Cultivation						
Intercropped †	0.197**	0.578***	0.326***	0.140*	0.198***	0.222***
	(0.077)	(0.097)	(0.084)	(0.078)	(0.077)	(0.080)
Share of Plot Area Under Improved Seeds	-0.010	-0.081	-0.036	-0.027	-0.013	-0.038
	(0.090)	(0.091)	(0.090)	(0.089)	(0.091)	(0.091)
Share of Plot Area Under Export Crops	0.055	0.025	-0.008	0.079	0.085	0.045
	(0.145)	(0.155)	(0.147)	(0.149)	(0.142)	(0.148)
Household Characteristics						
Household Size	0.042**	0.027*	0.035**	0.039**	-0.064	0.043**
	(0.017)	(0.017)	(0.017)	(0.017)	(0.054)	(0.017)
Child Dependency Ratio	-0.067	-0.085*	-0.086*	-0.074	-0.004	-0.067
	(0.052)	(0.051)	(0.051)	(0.052)	(0.064)	(0.053)
Agricultural Extension Receipt †	0.163**	0.067	0.157**	0.140*	0.164**	0.177**
	(0.072)	(0.074)	(0.074)	(0.071)	(0.072)	(0.074)
Access to Non-Farm Labor Income †	-0.267***	-0.239***	-0.240***	-0.260***	-0.269***	-0.281***
	(0.075)	(0.074)	(0.075)	(0.074)	(0.075)	(0.077)
Access to Non-Farm Non-Labor Income †	-0.208**	-0.120	-0.152*	-0.204**	-0.231***	-0.202**
	(0.088)	(0.089)	(0.088)	(0.087)	(0.088)	(0.088)
Wealth Index	0.044***	0.058***	0.051***	0.043***	0.036**	0.044***
	(0.015)	(0.016)	(0.016)	(0.015)	(0.015)	(0.016)
Agricultural Implement Access Index	0.086***	0.075**	0.076**	0.088***	0.084***	0.090***
	(0.029)	(0.030)	(0.030)	(0.029)	(0.029)	(0.030)
Distance to Nearest ADMARC (KM)	-0.014**	-0.008	-0.009	-0.015**	-0.013*	-0.014*
	(0.007)	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)
Household Agro-Ecological Zone Classification						
Tropic-warm/semiarid †	0.443***	-0.164	0.352**	0.367**	0.445***	0.451***
	(0.163)	(0.225)	(0.177)	(0.158)	(0.165)	(0.172)
Tropic-warm/subhumid †	0.021	0.114	0.209	-0.060	0.037	0.043
	(0.170)	(0.227)	(0.186)	(0.167)	(0.171)	(0.176)
Tropic-cool/semiarid †	-0.029	-0.292	-0.094	-0.044	-0.050	0.008
	(0.181)	(0.267)	(0.198)	(0.176)	(0.184)	(0.189)
Observations	4,343	4,343	4,261	4,266	4,343	4,314
Adjusted R-Squared	0.103	0.135	0.120	0.102	0.105	0.102

Table H.14: Exploring the Presence of Omitted Variable Bias in Base OLS Regression Results Underlying the RIF Decomposition at the 50th

Dependent Variable: RIF(Log[Plot Value of Output (MK)/HA])

Pooled Sample

		Category of Additional Covariates Integrated into the Base Regression				
	Base	*District Fixed Effects*	*Plot Geospatial Characteristics*	*Other Plot Characteristics*	*Household Characteristics*	*Community Characteristics*
Plot Manager Characteristics						
Female †	-0.075**	-0.086***	-0.074**	-0.079***	-0.080***	-0.071**
	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)
Manager & Owner Overlap †	0.019	0.007	0.025	0.018	0.018	0.021
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
Age (Years)	-0.000	-0.000	-0.000	-0.000	0.002	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Years of Schooling	0.009***	0.006*	0.008***	0.008***	0.008***	0.008***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Plot Area						
Log[GPS-Based Plot Area (HA)]	-0.277***	-0.288***	-0.279***	-0.279***	-0.273***	-0.280***
	(0.025)	(0.025)	(0.026)	(0.025)	(0.025)	(0.025)
Log[GPS-Based Plot Area (HA) Squared]	0.032***	0.033***	0.032***	0.031***	0.033***	0.032***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Plot Non-Labor Input Use						
Incidence of Pesticide/Herbicide Use †	0.348***	0.366***	0.328***	0.347***	0.348***	0.341***
	(0.064)	(0.062)	(0.064)	(0.064)	(0.064)	(0.064)
Incidence of Organic Fertilizer Use †	0.051*	0.047	0.054*	0.049*	0.050*	0.046
	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)
Log[Inorganic Fertilizer Use (KG)/HA]	0.075***	0.080***	0.078***	0.075***	0.075***	0.076***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Plot Labor Input Use						
Log[Household Male Labor Use (Hours)/HA]	0.025***	0.024***	0.026***	0.025***	0.023***	0.026***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)
Log[Household Female Labor Use (Hours)/HA]	0.030***	0.028***	0.029***	0.031***	0.033***	0.031***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Log[Household Child Labor Use (Hours)/HA]	-0.002	-0.002	0.000	-0.001	-0.004	-0.003
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)
Log[Hired Labor Use (Days)/HA]	0.076***	0.071***	0.072***	0.077***	0.076***	0.074***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Log[Exchange Labor Use (Days)/HA]	0.047***	0.033**	0.041***	0.046***	0.047***	0.049***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.014)	(0.013)
Plot Location						
Elevation (M)	0.000***	0.000	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Distance to Household (KM)	-0.002***	-0.002**	-0.002**	-0.002***	-0.002***	-0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)

Table H.14 (Cont'd)

	Base	*Category of Additional Covariates Integrated into the Base Regression* District Fixed Effects	Plot Geospatial Characteristics	Other Plot Characteristics	Household Characteristics	Community Characteristics
Plot Cultivation						
Intercropped †	0.106***	0.277***	0.165***	0.096***	0.108***	0.112***
	(0.022)	(0.026)	(0.024)	(0.023)	(0.022)	(0.023)
Distance to Household (KM)	-0.002***	-0.002**	-0.002**	-0.002***	-0.002***	-0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Plot Cultivation						
Intercropped †	0.106***	0.277***	0.165***	0.096***	0.108***	0.112***
	(0.022)	(0.026)	(0.024)	(0.023)	(0.022)	(0.023)
Share of Plot Area Under Improved Seeds	0.121***	0.107***	0.121***	0.112***	0.119***	0.120***
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
Share of Plot Area Under Export Crops	0.905***	0.878***	0.907***	0.904***	0.906***	0.900***
	(0.029)	(0.030)	(0.030)	(0.031)	(0.029)	(0.029)
Household Characteristics						
Household Size	0.007	0.004	0.003	0.006	-0.061**	0.007
	(0.005)	(0.005)	(0.005)	(0.005)	(0.026)	(0.005)
Child Dependency Ratio	0.005	0.002	-0.002	0.005	0.008	0.006
	(0.017)	(0.017)	(0.017)	(0.017)	(0.023)	(0.017)
Agricultural Extension Receipt †	0.068***	0.034*	0.067***	0.065***	0.069***	0.082***
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
Access to Non-Farm Labor Income †	-0.060***	-0.065***	-0.059***	-0.057***	-0.064***	-0.057***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
Access to Non-Farm Non-Labor Income †	-0.044*	-0.014	-0.023	-0.047**	-0.045*	-0.047**
	(0.024)	(0.024)	(0.024)	(0.023)	(0.024)	(0.024)
Wealth Index	0.061***	0.064***	0.061***	0.061***	0.061***	0.061***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Agricultural Implement Access Index	0.045***	0.033***	0.043***	0.045***	0.045***	0.046***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Distance to Nearest ADMARC (KM)	0.004**	0.006***	0.004**	0.004**	0.004**	0.003*
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Household Agro-Ecological Zone Classification						
Tropic-warm/semiarid †	0.181***	0.000	0.202***	0.183***	0.184***	0.189***
	(0.045)	(0.062)	(0.049)	(0.045)	(0.046)	(0.046)
Tropic-warm/subhumid †	0.117**	0.114*	0.162***	0.114**	0.121**	0.132***
	(0.047)	(0.060)	(0.049)	(0.047)	(0.047)	(0.048)
Tropic-cool/semiarid †	0.090**	-0.057	0.109**	0.080*	0.085*	0.113**
	(0.046)	(0.059)	(0.049)	(0.046)	(0.047)	(0.047)
Observations	16,372	16,372	16,016	16,153	16,372	16,234
Adjusted R-Squared	0.204	0.224	0.219	0.203	0.205	0.205

Table H.15: Exploring the Presence of Omitted Variable Bias in Base OLS Regression Results Underlying the RIF Decomposition at the 50th

Dependent Variable: RIF(Log[Plot Value of Output (MK)/HA])

Male-Managed Plot Sample

		Category of Additional Covariates Integrated into the Base Regression				
	Base	*District Fixed Effects*	*Plot Geospatial Characteristics*	*Other Plot Characteristics*	*Household Characteristics*	*Community Characteristics*
Plot Manager Characteristics						
Female †						
Manager & Owner Overlap †	0.012	-0.011	0.015	0.010	0.013	0.015
	(0.023)	(0.024)	(0.023)	(0.023)	(0.023)	(0.023)
Age (Years)	-0.001	-0.001	-0.001	-0.001	0.001	-0.001*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Years of Schooling	0.009***	0.005	0.008**	0.008***	0.009***	0.009***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Plot Area						
Log[GPS-Based Plot Area (HA)]	-0.247***	-0.255***	-0.250***	-0.247***	-0.247***	-0.250***
	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)
Log[GPS-Based Plot Area (HA) Squared]	0.039***	0.041***	0.039***	0.038***	0.039***	0.039***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Plot Non-Labor Input Use						
Incidence of Pesticide/Herbicide Use †	0.305***	0.332***	0.266***	0.306***	0.312***	0.301***
	(0.069)	(0.067)	(0.069)	(0.069)	(0.070)	(0.069)
Incidence of Organic Fertilizer Use †	0.059*	0.045	0.067**	0.053	0.060*	0.054
	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)
Log[Inorganic Fertilizer Use (KG)/HA]	0.077***	0.082***	0.080***	0.077***	0.077***	0.078***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Plot Labor Input Use						
Log[Household Male Labor Use (Hours)/HA]	0.067***	0.066***	0.065***	0.067***	0.065***	0.067***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Log[Household Female Labor Use (Hours)/HA]	0.010	0.008	0.009	0.010	0.014*	0.012
	(0.008)	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)
Log[Household Child Labor Use (Hours)/HA]	0.006	0.007	0.007	0.007	0.004	0.006
	(0.005)	(0.005)	(0.006)	(0.005)	(0.006)	(0.006)
Log[Hired Labor Use (Days)/HA]	0.075***	0.069***	0.070***	0.076***	0.075***	0.072***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Log[Exchange Labor Use (Days)/HA]	0.045***	0.032*	0.037**	0.046***	0.047***	0.042**
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Plot Location						
Elevation (M)	0.000***	0.000	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Distance to Household (KM)	-0.002	-0.001	-0.001	-0.002**	-0.002	-0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)

Table H.15 (Cont'd)

	Base	District Fixed Effects	Plot Geospatial Characteristics	Other Plot Characteristics	Household Characteristics	Community Characteristics
		Category of Additional Covariates Integrated into the Base Regression				
Plot Cultivation						
Intercropped †	0.121***	0.303***	0.185***	0.116***	0.121***	0.125***
	(0.027)	(0.031)	(0.028)	(0.028)	(0.027)	(0.028)
Distance to Household (KM)	-0.002	-0.001	-0.001	-0.002**	-0.002	-0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Plot Cultivation						
Intercropped †	0.121***	0.303***	0.185***	0.116***	0.121***	0.125***
	(0.027)	(0.031)	(0.028)	(0.028)	(0.027)	(0.028)
Share of Plot Area Under Improved Seeds	0.123***	0.109***	0.115***	0.113***	0.120***	0.127***
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
Share of Plot Area Under Export Crops	0.944***	0.920***	0.939***	0.930***	0.940***	0.933***
	(0.033)	(0.034)	(0.033)	(0.036)	(0.033)	(0.033)
Household Characteristics						
Household Size	0.008	0.006	0.006	0.009	-0.074**	0.010
	(0.006)	(0.006)	(0.006)	(0.006)	(0.031)	(0.006)
Child Dependency Ratio	0.031	0.025	0.020	0.029	0.019	0.027
	(0.023)	(0.023)	(0.023)	(0.023)	(0.034)	(0.023)
Agricultural Extension Receipt †	0.058**	0.033	0.051**	0.056**	0.060**	0.062***
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
Access to Non-Farm Labor Income †	-0.063***	-0.069***	-0.069***	-0.062***	-0.065***	-0.059***
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Access to Non-Farm Non-Labor Income †	-0.007	0.015	0.015	-0.008	-0.004	-0.011
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
Wealth Index	0.070***	0.074***	0.068***	0.071***	0.071***	0.069***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Agricultural Implement Access Index	0.041***	0.027***	0.037***	0.041***	0.041***	0.041***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Distance to Nearest ADMARC (KM)	0.004**	0.006**	0.003	0.004**	0.004**	0.004*
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Household Agro-Ecological Zone Classification						
Tropic-warm/semiarid †	0.185***	-0.010	0.195***	0.188***	0.187***	0.204***
	(0.051)	(0.069)	(0.055)	(0.051)	(0.051)	(0.052)
Tropic-warm/subhumid †	0.108**	0.071	0.143**	0.112**	0.109**	0.130**
	(0.053)	(0.067)	(0.056)	(0.054)	(0.054)	(0.054)
Tropic-cool/semiarid †	0.145***	-0.015	0.176***	0.129**	0.143***	0.173***
	(0.052)	(0.066)	(0.056)	(0.052)	(0.053)	(0.054)
Observations	12,029	12,029	11,755	11,887	12,029	11,920
Adjusted R-Squared	0.211	0.231	0.226	0.210	0.211	0.212

Table H.16: Exploring the Presence of Omitted Variable Bias in Base OLS Regression Results Underlying the RIF Decomposition at the 50th

Dependent Variable: RIF(Log[Plot Value of Output (MK)/HA])

Female-Managed Plot Sample

		Category of Additional Covariates Integrated into the Base Regression				
	Base	*District Fixed Effects*	*Plot Geospatial Characteristics*	*Other Plot Characteristics*	*Household Characteristics*	*Community Characteristics*
Plot Manager Characteristics						
Female †						
Manager & Owner Overlap †	-0.039	-0.000	-0.032	-0.029	-0.040	-0.032
	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)
Age (Years)	0.002*	0.001	0.001	0.002*	0.003*	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Years of Schooling	0.014**	0.012**	0.014**	0.014**	0.013**	0.012**
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Plot Area						
Log[GPS-Based Plot Area (HA)]	-0.349***	-0.365***	-0.358***	-0.351***	-0.342***	-0.359***
	(0.049)	(0.049)	(0.050)	(0.050)	(0.049)	(0.049)
Log[GPS-Based Plot Area (HA) Squared]	0.011	0.010	0.013	0.010	0.013	0.008
	(0.011)	(0.011)	(0.012)	(0.011)	(0.011)	(0.011)
Plot Non-Labor Input Use						
Incidence of Pesticide/Herbicide Use †	0.560***	0.564***	0.612***	0.550***	0.547***	0.576***
	(0.139)	(0.144)	(0.144)	(0.138)	(0.141)	(0.136)
Incidence of Organic Fertilizer Use †	0.050	0.056	0.049	0.045	0.047	0.040
	(0.052)	(0.051)	(0.052)	(0.052)	(0.052)	(0.052)
Log[Inorganic Fertilizer Use (KG)/HA]	0.066***	0.069***	0.069***	0.062***	0.065***	0.065***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Plot Labor Input Use						
Log[Household Male Labor Use (Hours)/HA]	0.010	0.009	0.009	0.011	0.005	0.010
	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)	(0.007)
Log[Household Female Labor Use (Hours)/HA]	0.041**	0.036**	0.036**	0.043**	0.040**	0.041**
	(0.017)	(0.016)	(0.016)	(0.017)	(0.017)	(0.017)
Log[Household Child Labor Use (Hours)/HA]	-0.013*	-0.014**	-0.011	-0.012*	-0.013*	-0.013*
	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)	(0.007)
Log[Hired Labor Use (Days)/HA]	0.075***	0.074***	0.070***	0.078***	0.075***	0.076***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Log[Exchange Labor Use (Days)/HA]	0.031	0.023	0.018	0.027	0.027	0.029
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
Plot Location						
Elevation (M)	0.000**	0.000	0.000***	0.000**	0.000**	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Distance to Household (KM)	-0.003***	-0.003***	-0.004***	-0.003***	-0.003***	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)

Table H.16 (Cont'd)

		Category of Additional Covariates Integrated into the Base Regression				
	Base	*District Fixed Effects*	*Plot Geospatial Characteristics*	*Other Plot Characteristics*	*Household Characteristics*	*Community Characteristics*
Plot Cultivation						
Intercropped †	0.160***	0.311***	0.224***	0.152***	0.164***	0.172***
	(0.038)	(0.044)	(0.041)	(0.039)	(0.039)	(0.039)
Distance to Household (KM)	-0.003***	-0.003***	-0.004***	-0.003***	-0.003***	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Plot Cultivation						
Intercropped †	0.160***	0.311***	0.224***	0.152***	0.164***	0.172***
	(0.038)	(0.044)	(0.041)	(0.039)	(0.039)	(0.039)
Share of Plot Area Under Improved Seeds	0.133***	0.116***	0.152***	0.139***	0.132***	0.125***
	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)
Share of Plot Area Under Export Crops	0.864***	0.810***	0.876***	0.871***	0.889***	0.854***
	(0.070)	(0.073)	(0.072)	(0.073)	(0.069)	(0.070)
Household Characteristics						
Household Size	0.025***	0.022**	0.021**	0.022**	-0.014	0.024***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.025)	(0.009)
Child Dependency Ratio	-0.040	-0.038	-0.048*	-0.032	-0.010	-0.037
	(0.025)	(0.025)	(0.025)	(0.025)	(0.032)	(0.026)
Agricultural Extension Receipt †	0.188***	0.151***	0.183***	0.182***	0.188***	0.202***
	(0.039)	(0.040)	(0.040)	(0.040)	(0.039)	(0.040)
Access to Non-Farm Labor Income †	-0.022	-0.025	-0.011	-0.033	-0.024	-0.031
	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)
Access to Non-Farm Non-Labor Income †	-0.137***	-0.100**	-0.126***	-0.137***	-0.146***	-0.142***
	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)
Wealth Index	0.045***	0.048***	0.045***	0.044***	0.040***	0.042***
	(0.010)	(0.011)	(0.011)	(0.011)	(0.010)	(0.010)
Agricultural Implement Access Index	0.046***	0.033**	0.050***	0.045***	0.045***	0.045***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Distance to Nearest ADMARC (KM)	-0.002	0.001	0.003	-0.002	-0.001	-0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Household Agro-Ecological Zone Classification						
Tropic-warm/semiarid †	0.036	-0.148	0.091	0.009	0.030	0.010
	(0.094)	(0.138)	(0.100)	(0.094)	(0.096)	(0.096)
Tropic-warm/subhumid †	-0.029	0.005	0.082	-0.062	-0.022	-0.026
	(0.096)	(0.137)	(0.101)	(0.096)	(0.097)	(0.097)
Tropic-cool/semiarid †	-0.102	-0.289**	-0.138	-0.123	-0.118	-0.101
	(0.096)	(0.134)	(0.101)	(0.096)	(0.098)	(0.098)
Observations	4,343	4,343	4,261	4,266	4,343	4,314
Adjusted R-Squared	0.188	0.210	0.206	0.185	0.192	0.191

Table H.17: Exploring the Presence of Omitted Variable Bias in Base OLS Regression Results Underlying the RIF Decomposition at the 90th

Dependent Variable: RIF(Log[Plot Value of Output (MK)/HA])

Pooled Sample

		Category of Additional Covariates Integrated into the Base Regression				
	Base	*District Fixed Effects*	*Plot Geospatial Characteristics*	*Other Plot Characteristics*	*Household Characteristics*	*Community Characteristics*
Plot Manager Characteristics						
Female †	-0.082*	-0.097**	-0.077*	-0.088**	-0.083*	-0.080*
	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)
Manager & Owner Overlap †	-0.026	-0.018	-0.009	-0.027	-0.025	-0.029
	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)
Age (Years)	-0.002	-0.002	-0.002	-0.002	0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Years of Schooling	0.001	-0.002	0.001	0.001	0.002	0.002
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Plot Area						
Log[GPS-Based Plot Area (HA)]	0.065	0.055	0.064	0.057	0.069	0.070
	(0.054)	(0.054)	(0.055)	(0.055)	(0.054)	(0.054)
Log[GPS-Based Plot Area (HA) Squared]	0.188***	0.190***	0.190***	0.188***	0.188***	0.191***
	(0.021)	(0.020)	(0.021)	(0.021)	(0.020)	(0.021)
Plot Non-Labor Input Use						
Incidence of Pesticide/Herbicide Use †	0.703***	0.712***	0.705***	0.712***	0.711***	0.706***
	(0.137)	(0.137)	(0.138)	(0.141)	(0.138)	(0.138)
Incidence of Organic Fertilizer Use †	0.077	0.081	0.090	0.068	0.076	0.076
	(0.055)	(0.054)	(0.055)	(0.056)	(0.054)	(0.055)
Log[Inorganic Fertilizer Use (KG)/HA]	0.035***	0.039***	0.038***	0.037***	0.035***	0.036***
	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)	(0.007)
Plot Labor Input Use						
Log[Household Male Labor Use (Hours)/HA]	0.020**	0.019**	0.021**	0.019**	0.023**	0.023**
	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)	(0.009)
Log[Household Female Labor Use (Hours)/HA]	0.031**	0.028**	0.031**	0.032**	0.035**	0.035***
	(0.014)	(0.013)	(0.014)	(0.014)	(0.014)	(0.014)
Log[Household Child Labor Use (Hours)/HA]	0.000	-0.000	0.002	0.001	-0.007	-0.001
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Log[Hired Labor Use (Days)/HA]	0.104***	0.102***	0.103***	0.103***	0.104***	0.104***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Log[Exchange Labor Use (Days)/HA]	0.013	-0.001	0.007	0.013	0.014	0.013
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
Plot Location						
Elevation (M)	-0.000**	-0.000	-0.000	-0.000***	-0.000**	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Distance to Household (KM)	0.001	0.001	0.001	0.001	0.001	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)

Table H.17 (Cont'd)

		Category of Additional Covariates Integrated into the Base Regression				
	Base	*District Fixed Effects*	*Plot Geospatial Characteristics*	*Other Plot Characteristics*	*Household Characteristics*	*Community Characteristics*
Plot Cultivation						
Intercropped †	0.135***	0.239***	0.193***	0.120***	0.132***	0.135***
	(0.034)	(0.042)	(0.036)	(0.035)	(0.034)	(0.035)
Distance to Household (KM)	0.001	0.001	0.001	0.001	0.001	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Plot Cultivation						
Intercropped †	0.135***	0.239***	0.193***	0.120***	0.132***	0.135***
	(0.034)	(0.042)	(0.036)	(0.035)	(0.034)	(0.035)
Share of Plot Area Under Improved Seeds	0.023	0.018	0.033	0.016	0.022	0.026
	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)
Share of Plot Area Under Export Crops	3.062***	3.044***	3.066***	3.118***	3.057***	3.068***
	(0.107)	(0.107)	(0.108)	(0.111)	(0.108)	(0.108)
Household Characteristics						
Household Size	0.006	0.006	0.005	0.006	-0.118***	0.006
	(0.008)	(0.008)	(0.008)	(0.008)	(0.037)	(0.008)
Child Dependency Ratio	-0.006	-0.009	-0.015	-0.001	-0.018	-0.005
	(0.026)	(0.026)	(0.026)	(0.027)	(0.033)	(0.026)
Agricultural Extension Receipt †	0.052	0.020	0.049	0.055	0.055	0.057
	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.035)
Access to Non-Farm Labor Income †	-0.130***	-0.129***	-0.128***	-0.136***	-0.131***	-0.127***
	(0.031)	(0.032)	(0.031)	(0.032)	(0.031)	(0.032)
Access to Non-Farm Non-Labor Income †	-0.016	0.024	0.001	-0.025	-0.019	-0.016
	(0.039)	(0.039)	(0.040)	(0.041)	(0.039)	(0.040)
Wealth Index	0.060***	0.064***	0.061***	0.061***	0.061***	0.063***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Agricultural Implement Access Index	0.036***	0.022*	0.036***	0.035***	0.036***	0.034***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Distance to Nearest ADMARC (KM)	0.003	0.004	0.006**	0.002	0.003	0.003
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Household Agro-Ecological Zone Classification						
Tropic-warm/semiarid †	-0.200***	-0.206**	-0.160**	-0.190***	-0.208***	-0.180**
	(0.072)	(0.102)	(0.079)	(0.074)	(0.073)	(0.076)
Tropic-warm/subhumid †	-0.146*	0.009	-0.115	-0.140*	-0.151**	-0.131*
	(0.076)	(0.101)	(0.082)	(0.077)	(0.076)	(0.077)
Tropic-cool/semiarid †	-0.114	-0.231**	-0.080	-0.100	-0.122*	-0.107
	(0.073)	(0.092)	(0.077)	(0.074)	(0.074)	(0.076)
Observations	16,372	16,372	16,016	16,153	16,372	16,234
Adjusted R-Squared	0.258	0.265	0.267	0.258	0.259	0.259

Table H.18: Exploring the Presence of Omitted Variable Bias in Base OLS Regression Results Underlying the RIF Decomposition at the 90th

Dependent Variable: RIF(Log[Plot Value of Output (MK)/HA])

Male-Managed Plot Sample

		Category of Additional Covariates Integrated into the Base Regression				
	Base	*District Fixed Effects*	*Plot Geospatial Characteristics*	*Other Plot Characteristics*	*Household Characteristics*	*Community Characteristics*
Plot Manager Characteristics						
Female †						
Manager & Owner Overlap †	-0.021	-0.020	-0.003	-0.017	-0.017	-0.030
	(0.036)	(0.037)	(0.037)	(0.036)	(0.036)	(0.036)
Age (Years)	-0.001	-0.001	-0.001	-0.001	0.002	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Years of Schooling	-0.000	-0.003	-0.002	0.000	0.000	-0.001
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Plot Area						
Log[GPS-Based Plot Area (HA)]	0.101*	0.093	0.097	0.094	0.100*	0.103*
	(0.059)	(0.059)	(0.060)	(0.061)	(0.059)	(0.059)
Log[GPS-Based Plot Area (HA) Squared]	0.183***	0.186***	0.184***	0.185***	0.182***	0.184***
	(0.024)	(0.024)	(0.025)	(0.025)	(0.024)	(0.024)
Plot Non-Labor Input Use						
Incidence of Pesticide/Herbicide Use †	0.765***	0.767***	0.744***	0.703***	0.778***	0.768***
	(0.154)	(0.156)	(0.156)	(0.162)	(0.155)	(0.155)
Incidence of Organic Fertilizer Use †	0.127*	0.119*	0.142**	0.114*	0.123*	0.123*
	(0.065)	(0.065)	(0.066)	(0.067)	(0.065)	(0.066)
Log[Inorganic Fertilizer Use (KG)/HA]	0.036***	0.039***	0.039***	0.036***	0.035***	0.036***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Plot Labor Input Use						
Log[Household Male Labor Use (Hours)/HA]	0.084***	0.083***	0.078***	0.087***	0.082***	0.091***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Log[Household Female Labor Use (Hours)/HA]	0.005	0.003	0.007	0.008	0.011	0.007
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Log[Household Child Labor Use (Hours)/HA]	0.002	0.003	0.004	0.004	-0.001	0.000
	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)
Log[Hired Labor Use (Days)/HA]	0.117***	0.115***	0.111***	0.120***	0.116***	0.118***
	(0.017)	(0.017)	(0.017)	(0.018)	(0.017)	(0.017)
Log[Exchange Labor Use (Days)/HA]	0.015	0.005	0.012	0.013	0.015	0.014
	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)
Plot Location						
Elevation (M)	-0.000**	-0.000	-0.000	-0.000***	-0.000**	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Distance to Household (KM)	0.002	0.003	0.003	0.002	0.002	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)

Table H.18 (Cont'd)

		Category of Additional Covariates Integrated into the Base Regression				
	Base	*District Fixed Effects*	*Plot Geospatial Characteristics*	*Other Plot Characteristics*	*Household Characteristics*	*Community Characteristics*
Plot Cultivation						
Intercropped †	0.097**	0.196***	0.138***	0.074*	0.093**	0.092**
	(0.040)	(0.047)	(0.041)	(0.041)	(0.040)	(0.041)
Distance to Household (KM)	0.002	0.003	0.003	0.002	0.002	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Plot Cultivation						
Intercropped †	0.097**	0.196***	0.138***	0.074*	0.093**	0.092**
	(0.040)	(0.047)	(0.041)	(0.041)	(0.040)	(0.041)
Share of Plot Area Under Improved Seeds	-0.010	-0.014	-0.004	-0.020	-0.010	-0.006
	(0.034)	(0.034)	(0.034)	(0.035)	(0.034)	(0.034)
Share of Plot Area Under Export Crops	2.950***	2.933***	2.962***	2.972***	2.947***	2.933***
	(0.114)	(0.114)	(0.115)	(0.118)	(0.114)	(0.114)
Household Characteristics						
Household Size	0.005	0.006	0.004	0.004	-0.120***	0.005
	(0.010)	(0.010)	(0.010)	(0.010)	(0.045)	(0.010)
Child Dependency Ratio	0.029	0.024	0.023	0.036	0.038	0.026
	(0.037)	(0.037)	(0.037)	(0.037)	(0.053)	(0.037)
Agricultural Extension Receipt †	0.039	0.013	0.038	0.042	0.044	0.038
	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)
Access to Non-Farm Labor Income †	-0.116***	-0.112***	-0.126***	-0.124***	-0.117***	-0.111***
	(0.036)	(0.037)	(0.036)	(0.036)	(0.036)	(0.036)
Access to Non-Farm Non-Labor Income †	0.004	0.030	0.018	0.000	0.004	0.014
	(0.045)	(0.045)	(0.045)	(0.046)	(0.045)	(0.045)
Wealth Index	0.056***	0.062***	0.057***	0.056***	0.056***	0.058***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Agricultural Implement Access Index	0.033**	0.019	0.032**	0.031**	0.031**	0.030**
	(0.014)	(0.015)	(0.015)	(0.015)	(0.014)	(0.014)
Distance to Nearest ADMARC (KM)	0.007**	0.008**	0.008**	0.007**	0.007**	0.007**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Household Agro-Ecological Zone Classification						
Tropic-warm/semiarid †	-0.142*	-0.160	-0.096	-0.118	-0.156**	-0.120
	(0.075)	(0.104)	(0.084)	(0.076)	(0.076)	(0.079)
Tropic-warm/subhumid †	-0.062	0.054	-0.045	-0.043	-0.076	-0.048
	(0.079)	(0.105)	(0.087)	(0.080)	(0.079)	(0.081)
Tropic-cool/semiarid †	-0.011	-0.141	0.026	0.025	-0.024	0.008
	(0.076)	(0.094)	(0.081)	(0.077)	(0.076)	(0.080)
Observations	12,029	12,029	11,755	11,887	12,029	11,920
Adjusted R-Squared	0.271	0.278	0.278	0.271	0.272	0.272

Table H.19: Exploring the Presence of Omitted Variable Bias in Base OLS Regression Results Underlying the RIF Decomposition at the 90th

Dependent Variable: RIF(Log[Plot Value of Output (MK)/HA])

Female-Managed Plot Sample

		Category of Additional Covariates Integrated into the Base Regression				
	Base	*District Fixed Effects*	*Plot Geospatial Characteristics*	*Other Plot Characteristics*	*Household Characteristics*	*Community Characteristics*
Plot Manager Characteristics						
Female †						
Manager & Owner Overlap †	-0.088	-0.040	-0.096	-0.106	-0.098	-0.101
	(0.083)	(0.086)	(0.085)	(0.086)	(0.083)	(0.082)
Age (Years)	-0.002	-0.002	-0.001	-0.002	0.000	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Years of Schooling	0.019*	0.015	0.026**	0.021*	0.020*	0.017
	(0.011)	(0.012)	(0.012)	(0.012)	(0.011)	(0.011)
Plot Area						
Log[GPS-Based Plot Area (HA)]	0.035	-0.004	0.027	0.040	0.043	0.072
	(0.105)	(0.107)	(0.112)	(0.109)	(0.107)	(0.105)
Log[GPS-Based Plot Area (HA) Squared]	0.213***	0.211***	0.220***	0.217***	0.216***	0.221***
	(0.033)	(0.033)	(0.035)	(0.034)	(0.033)	(0.033)
Plot Non-Labor Input Use						
Incidence of Pesticide/Herbicide Use †	0.232	0.223	0.254	0.229	0.236	0.242
	(0.320)	(0.313)	(0.334)	(0.328)	(0.316)	(0.322)
Incidence of Organic Fertilizer Use †	-0.038	0.001	-0.058	-0.032	-0.038	-0.045
	(0.094)	(0.095)	(0.097)	(0.097)	(0.093)	(0.094)
Log[Inorganic Fertilizer Use (KG)/HA]	0.023	0.027*	0.029**	0.022	0.022	0.024*
	(0.014)	(0.014)	(0.014)	(0.015)	(0.014)	(0.014)
Plot Labor Input Use						
Log[Household Male Labor Use (Hours)/HA]	-0.034***	-0.036***	-0.035***	-0.035**	-0.030*	-0.032**
	(0.013)	(0.013)	(0.013)	(0.014)	(0.017)	(0.013)
Log[Household Female Labor Use (Hours)/HA]	0.067**	0.061*	0.061*	0.069**	0.066**	0.079**
	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)
Log[Household Child Labor Use (Hours)/HA]	-0.001	-0.004	0.002	-0.001	-0.011	-0.002
	(0.013)	(0.013)	(0.014)	(0.014)	(0.015)	(0.013)
Log[Hired Labor Use (Days)/HA]	0.144***	0.149***	0.139***	0.142***	0.146***	0.148***
	(0.033)	(0.033)	(0.034)	(0.035)	(0.033)	(0.033)
Log[Exchange Labor Use (Days)/HA]	0.022	0.008	-0.003	0.024	0.020	0.019
	(0.039)	(0.040)	(0.038)	(0.040)	(0.039)	(0.039)
Plot Location						
Elevation (M)	0.000	-0.000	0.000*	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Distance to Household (KM)	-0.001	-0.000	-0.001	-0.002*	-0.000	-0.002**
	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)

Table H.19 (Cont'd)

		Category of Additional Covariates Integrated into the Base Regression				
	Base	*District Fixed Effects*	*Plot Geospatial Characteristics*	*Other Plot Characteristics*	*Household Characteristics*	*Community Characteristics*
Plot Cultivation						
Intercropped †	0.152**	0.264***	0.226***	0.135*	0.155**	0.132*
	(0.070)	(0.084)	(0.076)	(0.074)	(0.071)	(0.071)
Distance to Household (KM)	-0.001	-0.000	-0.001	-0.002*	-0.000	-0.002**
	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Plot Cultivation						
Intercropped †	0.152**	0.264***	0.226***	0.135*	0.155**	0.132*
	(0.070)	(0.084)	(0.076)	(0.074)	(0.071)	(0.071)
Share of Plot Area Under Improved Seeds	-0.022	-0.058	0.019	-0.021	-0.026	-0.014
	(0.073)	(0.074)	(0.074)	(0.075)	(0.074)	(0.074)
Share of Plot Area Under Export Crops	3.521***	3.426***	3.616***	3.554***	3.530***	3.518***
	(0.336)	(0.338)	(0.344)	(0.345)	(0.334)	(0.334)
Household Characteristics						
Household Size	0.054***	0.053***	0.052***	0.053***	-0.029	0.051***
	(0.017)	(0.017)	(0.017)	(0.018)	(0.045)	(0.017)
Child Dependency Ratio	-0.121***	-0.111***	-0.150***	-0.111**	-0.087	-0.115***
	(0.042)	(0.042)	(0.042)	(0.044)	(0.053)	(0.042)
Agricultural Extension Receipt †	0.147*	0.089	0.121	0.147*	0.147*	0.169**
	(0.076)	(0.079)	(0.076)	(0.078)	(0.076)	(0.077)
Access to Non-Farm Labor Income †	-0.049	-0.046	-0.018	-0.050	-0.054	-0.048
	(0.067)	(0.067)	(0.070)	(0.069)	(0.067)	(0.066)
Access to Non-Farm Non-Labor Income †	-0.043	0.009	-0.008	-0.073	-0.069	-0.060
	(0.084)	(0.084)	(0.087)	(0.087)	(0.084)	(0.084)
Wealth Index	0.069***	0.082***	0.070***	0.069***	0.064***	0.059**
	(0.023)	(0.024)	(0.024)	(0.024)	(0.023)	(0.024)
Agricultural Implement Access Index	0.045*	0.023	0.049**	0.037	0.043*	0.034
	(0.025)	(0.026)	(0.024)	(0.026)	(0.025)	(0.025)
Distance to Nearest ADMARC (KM)	-0.011**	-0.009	-0.004	-0.012**	-0.011**	-0.009*
	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)
Household Agro-Ecological Zone Classification						
Tropic-warm/semiarid †	-0.185	-0.077	-0.089	-0.218	-0.156	-0.310
	(0.188)	(0.258)	(0.210)	(0.195)	(0.191)	(0.191)
Tropic-warm/subhumid †	-0.178	0.058	-0.094	-0.190	-0.146	-0.282
	(0.194)	(0.262)	(0.211)	(0.201)	(0.196)	(0.196)
Tropic-cool/semiarid †	-0.298	-0.317	-0.239	-0.326*	-0.291	-0.403**
	(0.190)	(0.245)	(0.202)	(0.195)	(0.195)	(0.192)
Observations	4,343	4,343	4,261	4,266	4,343	4,314
Adjusted R-Squared	0.204	0.209	0.217	0.200	0.206	0.208

REFERENCES

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). "The colonial origins of comparative development: An empirical investigation." *American Economic Review*, 91 (5), 1369-1401.

Alene, A. D., Manyong, V. M., Omanya, G. O., Mignouna, H. D., Bokanga, M., & Odhiambo, G. D. (2008). "Economic efficiency and supply response of women as farm managers: Comparative evidence from Western Kenya." *World Development* 36 (7), 1247–1260.

Akresh, R. (2005). "Understanding pareto inefficient intrahousehold allocations." Discussion Paper No. 1858, IZA. Retrieved from http://ftp.iza.org/dp1858.pdf.

Altonji, J. (1988). "The effects of family background and school characteristics on education and labor market outcomes." Manuscript, Department of Economics, Northwestern University.

Altonji, J., Elder, T., & Taber, C. (2005). "Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools." *Journal of Political Economy*, 113, 151-184.

Blinder, A. (1973). "Wage discrimination: reduced form and structural estimates." *Journal of Human Resources*, 8, 436–455.

Chen, S., & Ravallion, M. (2008). "The developing world is poorer than we thought but no less successful in the fight against poverty." Washington, DC: The World Bank.

Doepke, M., & Tertilt, M. (2011). "Does female empowerment promote economic development?" Discussion Paper No. DP8441, CEPR.

FAO (2011). T*he state of food and agriculture: Women in agriculture – closing the gender gap for development.* Rome, Italy: FAO. Retrieved from http://www.fao.org/docrep/013/i2050e/i2050e.pdf.

_____ (2009). H*ow to feed the world in 2050. High level expert forum – the special challenge for sub-Saharan Africa.* Rome, Italy: FAO. Retrieved from http://www.fao.org/fileadmin/templates/wsfs/docs/Issues_papers/HLEF2050_Africa.pdf.

Firpo, S., Fortin, N. M., & Lemieux, T. (2009). "Unconditional quantile regressions." *Econometrica*, 77 (3), 953–973.

Fortin, N. M. (2006). "Greed, altruism, and the gender wage gap." Working Paper, Department of Economics, University of British Columbia. Retrieved from http://faculty.arts.ubc.ca/nfortin/Fortinat8.pdf.

Fortin, N., T. Lemieux, S. Firpo (2011). "Decomposition methods." In O. Ashenfelter & D. Card (Eds.), *Handbook of labor economics* (Vol. 4, Part A, 1-102) Amsterdam, Netherlands: North-Holland.

Gilbert, R. A., Sakala, W. D., & Benson, T. D. (2002). "Gender analysis of a nationwide cropping system trial survey in Malawi." *African Studies Quarterly*, 6 (1), Retrieved from http://web.africa.ufl.edu/asq/v6/v6i1a9.htm.

Goldstein, M., & Udry, C. (2008). "The profits of power: Land rights and agricultural investment in Ghana." *Journal of Political Economy* 116 (6), 981–1022.

Holden, S.T., R. Lunduka (2010). "Too poor to be efficient? Impacts of the targeted fertilizer subsidy program in Malawi on farm plot level input use, crop choice and land productivity." Noragric Report No. 55, Department of Economics and Resource Management, Norwegian University of Life Sciences.

Irz, X., Lin, L., Thirtle, C., & Wiggins, S. (2001). “Agricultural productivity growth and poverty alleviation.” *Development Policy Review*, 19 (4), 449-466.

Imbens, W.G., & Rubin, B.D. (2009). *Causal inference in statistics, and in the social and biomedical sciences*. New York, NY: Cambridge University Press.

Jann, B. (2008). “The Blinder-Oaxaca decomposition for linear regression models.” *The Stata Journal*¸ 8 (4), 453-479.

Larson, D. F., Otsuka, K., Matsumoto, T., & Kilic, T. (2012). “Should African rural development strategies depend on smallholder farms? An exploration of the inverse productivity hypothesis.” Policy Research Working Paper No. 6190, The World Bank. Retrieved from http://elibrary.worldbank.org/content/workingpaper/10.1596/1813-9450-6190.

Ligon, E., & Sadoulet, E. (2008). “Estimating the effects of aggregate agricultural growth on the distribution of expenditures.” Background Paper for the World Bank World Development Report 2008.

Malawi Ministry of Agriculture and Food Security (MoAFS) (2009). *The 2009/10 farm input subsidy programme implementation guidelines.* Lilongwe, Malawi: MoAFS.

Malawi National Statistical Office (NSO) (2012). *Malawi third integrated household survey 2010-2011 basic information document.* Zomba, Malawi: NSO.

Moock, P. R. (1976). “The efficiency of women as farm managers: Kenya.” *American Journal of Agricultural Economics,* 58 (5), 831–835.

Murphy, K. M., & Topel, R. H. (1990). “Efficiency wages reconsidered: Theory and evidence.” In Y. Weiss & R. H. Topel (Eds.), *Advances in the theory and measurement of unemployment* (204-240). New York, NY: St. Martin’s Press.

Neumark, D. (1988). “Employers’ discriminatory behavior and the estimation of wage discrimination.” *Journal of Human Resources*, 23, 279–295.

Oaxaca, R. (1973). “Male-female wage differentials in urban labor markets.” *International Economic Review*, 14, 693–709.

Oladeebo, J. O., & Fajuyigbe, A. A. (2007). “Technical efficiency of men and women upland rice farmers in Osun State, Nigeria.” *Journal of Human Ecology*, 22 (2), 93–100.

O’Neill, J. E., & O’Neill, D. M. (2006). “What do wage differentials tell about labor market discrimination?” In S. W. Polachek, C. Chiswick & H. Rapoport (Eds.) *The economics of immigration and social diversity* (Research in labor economics series, Vol. 24, 293-357). Amsterdam, Netherlands: Elsevier.

Peterman, A., Quisumbing, A., Behrman, J., & Nkonya, E. (2011). “Understanding the complexities surrounding gender differences in agricultural productivity in Nigeria and Uganda.” *Journal of Development Studies*, 47 (10), 1482-1509.

Quisumbing, A., Payongayong, E., Aidoo, J. B., & Otsuka, K. (2001). “Women’s land rights in the transition to individualized ownership: Implications for the management of tree resources in western Ghana.” *Economic Development and Cultural Change*, 50 (1), 157–182.

Ricker-Gilbert, J., & Jayne, T. (2012). “Do fertilizer subsidies boost staple crop production and reduce poverty across the distribution of smallholders in Africa? Quantile regression results from Malawi.” Selected paper prepared for presentation at the International Association of Agricultural Economists (IAAE) Triennial Conference, Foz do Iguaçu, Brazil, 18-24 August, 2012. Retrieved from http://iaae.confex.com/iaae/iaae28/webprogram/Paper16263.html.

________________ (2011). “What are the enduring effects of fertilizer subsidy programs on recipient farm households? Evidence from Malawi.” Staff Paper, Department of Agricultural, Food and Resource Economics, Michigan State University.

Saito, K. A., Mekonnen, H., & Spurling, D. (1994). “Raising the productivity of women farmers in sub-Saharan Africa.” Discussion Paper Series 230, Africa Technical Department, The World Bank. Retrieved from http://elibrary.worldbank.org/content/book/9780821327494.

Schultz, T. P. (2001). “Women’s role in the agricultural household bargaining and human capital investments.” In B. Gardner & G. Rausser (Eds.), *Agricultural and resource economics handbook* (383–456). Amsterdam, Netherlands: Elsevier.

Theroux, P. (2004). *Dark star safari: Overland from Cairo to Cape Town.* New York, NY: Houghton Mifflin Company.

Tiruneh, A., Testfaye, T., Mwangi, W., & Verkuijl, H. (2001). “Gender differentials in agricultural production and decision-making among smallholders in Ada, Lume, and Gimbichu woredas of the central highlands of Ethiopia.” International Maize and Wheat Improvement Center and Ethiopian Agricultural Research Organization: Mexico, DF, and Addis Ababa. Retrieved from
http://repository.cimmyt.org/xmlui/bitstream/handle/10883/1018/73252.pdf?sequence=1

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

Vargas Hill, R., & Vigneri, M. (2011). “Mainstreaming gender sensitivity in cash crop markets supply chains.” Working Paper No. 11-08, ESA, FAO. Retrieved from http://www.fao.org/docrep/013/am313e/am313e00.pdf.

World Bank (WB) (2011). W*orld development report 2012: Gender equality and development.* Washington, DC: The World Bank. Retrieved from http://siteresources.worldbank.org/INTWDR2012/Resources/7778105-1299699968583/7786210-1315936222006/Complete-Report.pdf.

_________________ (2007). *Malawi: Country assistance strategy FY2007-FY2010.* Washington, DC: The World Bank.