Feed the Future Innovation Lab for Food Security Policy

Research Paper 158 December 2019

Zambia Buy-in

CAN AGRICULTURAL PRODUCTIVITY GROWTH SHAPE THE DEVELOPMENT OF THE NON-FARM RURAL ECONOMY? GEOGRAPHICALLY LOCALIZED EVIDENCE FROM ZAMBIA

By

Jason Snyder, Thomas Jayne, Jordan Chamberlin, Paul Samboko, Nicole Mason

Food Security Policy *Research Papers*

This *Research Paper* series is designed to timely disseminate research and policy analytical outputs generated by the USAID funded Feed the Future Innovation Lab for Food Security Policy (FSP) and its Associate Awards. The FSP project is managed by the Food Security Group (FSG) of the Department of Agricultural, Food, and Resource Economics (AFRE) at Michigan State University (MSU), and implemented in partnership with the International Food Policy Research Institute (IFPRI) and the University of Pretoria (UP). Together, the MSU-IFPRI-UP consortium works with governments, researchers and private sector stakeholders in Feed the Future focus countries in Africa and Asia to increase agricultural productivity, improve dietary diversity and build greater resilience to challenges like climate change that affect livelihoods.
The papers are aimed at researchers, policy makers, donor agencies, educators, and international development practitioners. Selected papers will be translated into French, Portuguese, or other languages.

Copies of all FSP Research Papers and Policy Briefs are freely downloadable in pdf format from the following Web site: https://www.canr.msu.edu/fsp/publications/
Copies of all FSP papers and briefs are also submitted to the USAID Development Experience Clearing House (DEC) at: http://dec.usaid.gov/

AUTHORS

Jason Snyder – Research Associate, Department of Agricultural, Food, and Resource Economics, Michigan State University

Thomas Jayne – MSU Foundation Professor, Department of Agricultural, Food, and Resource Economics, Michigan State University

Jordan Chamberlin – Spatial Economist, CIMMYT

Paul Samboko – PhD Student, Department of Agricultural, Food, and Resource Economics, Michigan State University

Nicole Mason – Associate Professor, Department of Agricultural, Food, and Resource Economics, Michigan State University

Authors' Acknowledgment:

This work was made possible by the generous support of the American People provided to the Feed the Future Innovation Lab for Food Security Policy (FSP) through the United States Agency for International Development (USAID) under Cooperative Agreement No. AID-OAA-L-13-00001 (Zambia Buy-In). This work was also supported by the US Department of Agriculture (USDA) National Institute of Food and Agriculture and Michigan AgBioResearch [project number MICL02501]. The contents are the sole responsibility of the authors and do not necessarily reflect the views of FSP, USAID, USDA, the United States Government, or Michigan AgBioResearch.

This study is made possible by the generous support of the American people provided to the Feed the Future Innovation Lab for Food Security Policy (FSP) through the United States Agency for International Development (USAID) under Cooperative Agreement No. AID-OAA-L-13-00001 (Zambia Buy-In). The contents are the sole responsibility of the authors and do not necessarily reflect the views of USAID or the United States Government.

Copyright © 2019, Michigan State University and IAPRI. All rights reserved. This material may be reproduced for personal and not-for-profit use without permission from but with acknowledgement to MSU and IAPRI.

Published by the Department of Agricultural, Food, and Resource Economics, Michigan State University, Justin S. Morrill Hall of Agriculture, 446 West Circle Dr., Room 202, East Lansing,

ABSTRACT

There is a general consensus in the economics literature that growth in agricultural productivity is an engine of economic structural transformation in low income countries via indirect linkages and multiplier effects. However, the empirical support for this hypothesis in the Africa is very limited, and criticisms have been raised as to its applicability in the African context. In this study, we estimate the strength of possible 'labor linkages' among small farmers in Zambia, helping to provide a much-needed empirical micro-foundation in the African context. In particular, we use nationally representative surveys to estimate the relationship between multiple years of lagged district level crop productivity and small farm household non-farm labor participation. We find that a doubling of average district crop productivity leads to a 13%-17% increase in non-farm labor participation among farm households. Moreover, this effect is most pronounced among smaller farms; a doubling of median district crop productivity among farms under two hectares cultivated leads to a 24%-31% increase in in non-farm labor participation among non-farm households. The results lend some credibility to the structural transformation hypothesis, and in particular, the idea of labor linkages, in the African setting

Keywords: Agricultural productivity, non-farm labor, Zambia.

JEL codes: Q12, Q18

TABLE OF CONTENTS

LIST OF TABLES

ACRONYMS

AE	Adult Equivalent
AIC	Akaike Information Criterion
CFS	Crop Forecast Survey
CSO	Central Statistical Office
CV	Coefficient of Variation
FISP	Farmer Input Support Program
FRA	Food Reserve Agency
IAPRI	Indaba Agricultural Policy Research Institute
IFAD	International Fund for Agricultural Development
MoA	Ministry of Agriculture
RALS	Rural Agricultural Livelihoods Survey
TAMSAT	Tropical Applications of Meteorology using Satellite Data and Ground-based Observations
TLU	Tropical Livestock Unit
ZMW	Zambian Kwacha

1. INTRODUCTION

A longstanding stylized fact from the development economics literature is that agricultural growth is central to the structural transformation process in developing economies. This literature dates back to Lewis (1954), Johnston and Mellor (1961), and Ranis and Fei (1961) and theorizes linkages between agricultural growth and the rest of the economy via consumption, factor markets and production, and wage effects (Hirschman 1958; Mellor 1976; Hirschman 1992; Block and Timmer 1994; Haggblade, Hazell, and Reardon 2007; Barrett, Carter, and Timmer 2010). A practical conclusion from this literature is that policies that discriminate against agriculture are generally counterproductive for structural transformation and for sustained economy-wide growth (Dennis and İşcan 2011)

However, emerging research has started to question this longstanding consensus about agricultural growth and structural transformation, especially in the context of rapidly changing structural dynamics in sub-Saharan Africa (Ellis 2005; Diao et al. 2012; Hazell 2013; Collier and Dercon 2014; Dercon and Gollin 2014). Some have argued that either cities are the main drivers of rural agricultural growth (Jacobs 2016), or that the rural non-farm sectors in secondary cities are the main drivers of inclusive economic growth (Christiaensen and Todo 2013). Ultimately, the role of agricultural growth as a driver of structural transformation is an empirical question. However, there is not much micro-level empirical evidence for how the structural transformation model plays out in sub-Saharan Africa. Part of the reason for the relative dearth of literature on this topic is that data on agricultural productivity over time has been mostly unavailable. Djoumessi, Kamdem, and Nembot (2019) estimate the effect of multi-year lagged productivity growth on non-farm employment across 13 African countries over several years using international databases, but African-based studies that estimate this effect at the micro-level are rare. A recent exception is the work of Takeshima, Amare, and Mavrotas (2018), who, using a three-wave panel LSMS household dataset from Nigeria, show that agricultural productivity is positively associated with non-farm capital and labor activities within the household. However, most analyses focus on the transition from farm to non-farm households but are not able to capture the early micro-level change in the farm to non-farm labor ratio within farm households themselves.

In this study, we estimate the impacts of district-level multi-year lagged crop productivity on non-farm labor participation, among farm households in rural Zambia. We test the hypothesis that localized farm productivity indirectly stimulates farm household off-farm employment, which would support the view that agriculture matters. Our primary dataset is a nationally representative household panel survey, the Rural Agricultural Livelihoods Survey (RALS), with rounds collected in 2012 and 2015. This dataset is merged with multiple years of lagged productivity measures at the district level (from 2004/05 through 2012/13), derived from the annual Crop Forecast Surveys (CFS). There are 72 districts in Zambia and the CFS data is statistically representative at district level. Hence, we are able identify at quite localized levels the effect of lagged agricultural productivity on local employment. In addition to estimating the effect of localized (both median and mean) district lagged productivity on household non-farm labor participation, we also estimate the median productivity among farms that are less than 2 hectares, and greater than or equal to 2 hectares (included in the same model). The models include other household- and district-level covariates drawn from the RALS panel survey. This approach affords us the following benefits. First, we can test the robustness of the relationship to alternative district summary measures of productivity, including the mean and median, and the median for smaller and larger farms. Second, we can account for the effects over time using multiple lag years with both moving average and the flexible Almon lag specification. Third, we can control for several observed and unobserved household and regional level covariates, something that is lacking in cross-country analysis. Finally, we can evaluate the effects of agricultural productivity on the transition to non-farm labor within rural farm households, enabling us to capture the earliest signs of structural transformation. Overall, this study helps to provide, in addition to the recent work of Takeshima, Amare, and Mavrotas (2018), a much needed empirical and micro-level foundation on this topic in the context of sub-Saharan Africa.

The rest of this article is divided into the following sections. First, we outline the conceptual underpinnings of our analysis, reviewing the structural transformation literature in the context of more recent trends in Africa. Second, we present the empirical model, estimation strategy, and data. Finally, we present the results and robustness checks, and conclude with a summary of the main contributions to the literature, as well as relevant policy implications for inclusive economic development. Our study aligns with the predictions from the broader literature that agricultural productivity is positively and significantly associated with the movement of labor off the farm. However, unlike most previous literature, we distinguish this effect within rural farm households and find some tentative evidence that this is especially true for district productivity gains for smaller farms in particular, and at the lower end of the productivity distribution.

2. ANALYTICAL FRAMEWORK

2.1 Conceptual underpinnings

There is a rich theoretical literature on how agriculture in developing countries is connected to the broader economy through a series of "linkages" based on consumption, production, factor input and output markets, and wages (Hirschman 1958; Mellor and Lele 1973; Hirschman 1992; Block and Timmer 1994; Haggblade, Hazell, and Reardon 2007; Barrett, Carter, and Timmer 2010). On the consumption side, agricultural productivity growth is said to lead to higher farmer incomes, which boosts demand for non-farm goods and services in the regional economy. On the production side, this growth increases activity in the downstream and upstream segments of the food system, e.g., fertilizer, seed and pesticide suppliers, food processors, traders and wholesalers, and food retailers. In factor markets, rising labor productivity incentivizes labor movement off the farm, while production surpluses may get invested in non-farm sectors of the economy. Finally, increases in agricultural productivity will lower prices for consumers by lowering the price of food, thus boosting their real wages. Our focus is on the labor factor market linkage. In particular, we draw upon the dual economy model of Lewis (1954) and Fei and Ranis (1963). Djoumessi, Kamdem, and Nembot (2019), applying this model in the African context, demonstrate analytically that an increase in farm productivity should lead to a decrease in agricultural labor, and an increase in employment off of the farm. We will test these assumptions empirically.

While this broad narrative is well established in the literature, it is important to consider the specific context of sub-Saharan Africa. There is speculation that the impact of agricultural growth on structural transformation is declining as the economies diversify, or, it was always too constrained to serve as a growth engine to begin with (Collier and Dercon, 2014). Hazell (2013) argues that agricultural led development is not a sure thing; many countries are experiencing a decline in farm sizes which can have mixed impacts, and that it is important to differentiate among types of farms, in particular, commercial farms, subsistence farms, and transition farms. Work by Yeboah and Jayne (2018) and IFAD (2019) suggests that the share of labor force involved with the post-farm segments of the agri-food system is rapidly increasing, while labor on-farm is declining. This trend is expected to continue over the next few decades (Tschirley et al. 2015). In addition, Jane Jacobs is well known for rejecting the idea of agriculturally driven structural transformation altogether, instead arguing that cities are responsible for coordinating rural and agricultural growth (Jacobs 2016). The impact of agricultural growth on structural transformation, and, labor movement out of agriculture, is ultimately an empirical and context-dependent question.

Until recently, this type of analysis has been difficult in Sub-Saharan Africa, due to a lack of available high-resolution data. Also, in order to evaluate the role of agricultural growth in structural transformation, several identification issues need to be considered. These include:

The first is causality. We live in a complex world with multiple non-linear feedback loops, and so perhaps the most fundamental challenge in economic analysis is the assertion of causality. Many of the early studies on structural transformation relied on cross-sectional data and were limited in their ability to parse correlation and causality (Tsakaok and Gardner 2007), especially given that agricultural productivity and non-farm labor effects may reciprocate in mutual influence. It is, therefore, imperative to have an identification strategy that credibly isolates the specific impact of agricultural productivity on non-farm labor transition. Utilizing lagged independent variables can help address this.

Accounting for unobserved factors is another challenge with identification, and, as Tsakok and Gardner (2007) also point out, this problem is exacerbated in studies that rely on cross-country data that is highly aggregated and fails to account for specific regional effects. This problem can be partially overcome due to the availability of household panel-survey data in certain countries, which allows researchers to control for both observed and unobserved household and regional effects.

In addition to helping address the causality issue, lagged independent variables also help account for dynamic effects that occur over time, as it is highly unlikely that the impact of agricultural productivity on non-farm labor participation would occur over a single year.

We address each of these identification challenges in the following model.

2.2. Empirical Model

Our approach is to identify the impact of cropland productivity at the district level on non-farm labor participation among rural farm households:

$$NF_{hdpy}^{L} = \beta_0 + \sum_{l=1}^{L} V_{dp,y-l}^{P} \beta_1^{y-l} + \boldsymbol{x_{hdpy}}\boldsymbol{\beta_2} + \boldsymbol{a_{hdpy}}\boldsymbol{\beta_3} + \beta_4 c_{dpy} + \boldsymbol{r_{hdpy}}\boldsymbol{\beta_5} + u_h + \lambda_y + \theta_p + \omega_{p,y} + v_{hdpy}$$

where $h, d, p,$ and y index the household, district, province, and year, respectively; l indexes the lag; L is the total number of lags; NF_{hdpy}^{L} represents either (a) number of adult household members that are receiving non-farm (excluding own-farm and farm wage) income, or (b) number of adult equivalents receiving non-farm income; $V_{dp,y-l}^{P}$ for $l=1, 2, \ldots, L$ are the key lagged variables of interest, the value of crop productivity per hectare, summarized by district. The β_1^{y-l}'s are the main parameters to be estimated. We define the following control variables: $\boldsymbol{x_{hdpy}}$ is a set of household demographic characteristics and quasi-fixed factors, including the number of adults in the household (or adult equivalents depending on the respective dependent variable), maximum and average adult (>15 years of age) education level in the household, a dummy variable equal to 1 if the household head is female; total number of hectares cultivated, tropical livestock units (TLU), and value of farm equipment, and finally, a dummy variable equal to 1 if the household in that year is below the total-sample median asset level; $\boldsymbol{a_{hdpy}}$ indicates perceived distance to infrastructure services, including the nearest district town, feeder road, tarmac road, boma (marketplace), and agricultural dealer; c_{dpy} is cell phone density (average number of cell phones per adult equivalent) – meant to capture relatively local changes in communication technology that might be associated with crop productivity.

We also account for short- and long-term rainfall, indicated by the vector $\boldsymbol{r_{hdpy}}$. The first variable is the total growing season rainfall (November-March), which directly affects farm activity in year *y*. Second, in order to account for the long-term farm impacts of rainfall, we include the average prior 16-years of growing season rainfall. Third and fourth, we account for moisture stress during both the most recent growing season, and for the prior 16 seasons, specified as the number of 20-day periods (overlapping) with less than 40 mm of rainfall. These account for a potential decline in productivity that can occur if rainfall falls below minimum short-period thresholds, independent of total growing season rainfall. Finally, we account for the prior 16-year growing season coefficient of variation (CV), meant to account for the variability of rainfall over time.

Finally, u_h accounts for time-invariant unobserved household heterogeneity; λ_y is the year fixed effect; θ_p are province fixed effects; $\omega_{p,y}$ are the interactions of province and year fixed effects; and v_{hdpy} is the error term.

3. ESTIMATION STRATEGY

Our hypothesis is that longer-term cropland productivity is a driver of structural transformation. We want to test whether land productivity drives transitions in labor allocation among rural farm households towards the non-farm sector. Our interest therefore is in deriving unbiased estimates of the β_1^{y-l}'s. Towards this end, we employ a correlated random effects (CRE) Mundlak-Chamberlain model (Mundlak 1978; Chamberlain 1984). While we include the household time-averages of the control variables, we do not include them for the lagged measures of district crop productivity. This approach allows us to control for time-constant unobserved heterogeneity (u_i) associated with the observed control variables, but to also avoid the potential attenuation bias and transitory noise that may result from demeaning, differencing, or taking the time average of our lagged variables that necessarily overlap due to the panel nature of the dataset. McKinnish (2008) has a discussion about how differenced lagged instruments are generally weak, and we expect that a similar logic applies for differenced lagged variables.

In the interest of robustness, we estimate multiple district-level measures of land productivity, including median, mean, and median for farms less than 2 hectares, and greater than or equal to two hectares (also estimated in the same model). This allows us to get a strong sense of not only the average impact of district productivity and labor transition into the non-farm sector but also changes for relatively larger and smaller farms. The summary measures have an implicit weighting by the value of output and area of cultivation, which we obtain by summarizing the value of output and the land cultivated before dividing them by district.

Finally, as indicated in our empirical model, we account for the effect of multiple lag years of district productivity on non-farm labor. We test two specifications, (a) including an Almon lag specification of multiple lag years (Almon 1965), and (b) creating a moving average of multiple lag year variables. For (a), we use an Almon lag because it provides allows for a lag structure distribution that is both parsimonious and flexible and avoids problems with multicollinearity that can arise with multiple lags left in an unrestricted form. It also allows for the possibility of a non-linear impact of crop productivity on non-farm labor over multiple years, perhaps taking multiple years to reach its peak, and even being negative in certain years. To implement the Almon lag, we approximate β_1^{l} in equation (1) by a 2nd degree polynomial, of the form $\beta_1^{t-l} = \alpha_0 + \alpha_1 j + \alpha_2 j^2$ for $l = 0, 1, 2, \dots, L\text{-}1$ such that the α's are estimated parameters. The β_1^{l-1}'s are then recovered from the α's and the elasticities are calculated. We determine the optimal lag length by (*L)* that minimizes the Akaike information criterion (AIC) in each respective model (based on guidance by Pindyck and Rubinfeld 1997 and Gujarati 2003), while setting three lags as the minimum lag length. The lag length for each model in (a) is also used for each corresponding model in (b), i.e., the moving average models.

We believe that our estimation strategy is effective in addressing threats to internal validity. These include potential omitted effects, non-linear dynamic effects, and correlation instead of causality. The great advantage of our dataset is that it includes multiple years of crop productivity data that, while not a panel at the household level, can be summarized as a panel at the district level, allowing us to estimate the long run and dynamic impact on household non-farm labor. It also includes a highly detailed household-level panel dataset that allows us to control for both highly localized time-constant unobserved heterogeneity, and time-varying heterogeneity and is not possible for similar studies that rely mostly on regional and national level data. We do not explicitly account for policy and operational changes of subsidy programmes under the Farmer Input Support Programme (FISP) and Zambia Food Reserve Agency (FRA), both an essential source of time-varying effects, but we believe that their primary mechanism of influence is via district-lagged productivity.[5]

4. DATA

We use three main sources of data. First is the Rural Agricultural Livelihoods Survey (RALS), a two-wave, nationally representative panel survey of smallholder farming households, that covers the agricultural years of 2010/11 and 2013/14 (from October – September) and the respective marketing years of 2011/12 and 2012/13 (from May – April). The collection was performed by the Indaba Agricultural Policy Research Institute (IAPRI), collaborating with the Zambia Central Statistical Office (CSO) and the Ministry of Agriculture (MoA) in June through July of 2012 and 2015, respectively. It is from this that we derive the key dependent variables, i.e., number of adults and adult equivalents, as well as all the control variables included in **x**, **a**, and c. The survey has a two-stage sample design, with a probability proportional to size sample design. Standard enumeration areas (or "clusters") are the primary sampling unit, stratified by districts. Households are the secondary sampling unit, stratified into three household categories, based on land cultivated, livestock raised, and types of crops produced. For more details on how the household categories are defined and the survey and sampling design, see CSO (2012) and CSO et al. (2012), respectively. There were 8,839 households surveyed in RALS 2012, and 7,254 (82%) of these were successfully re-interviewed in RALS 2015. We use the Wooldridge (2010) regression approach in each of our models to test for attrition bias.[7]

In table 1 we show the distribution of hectares cultivated for the households in the RALS sample. In both years most households cultivated less than two hectares of land, however, we see an increase in the share of farmers cultivating slightly larger land in 2015. From 2012 to 2015, the percentage of households cultivating less than two hectares decreased from 72.4% to 61.1%, while the percentage cultivating 2-5 hectares increased from 23.6% to 31.8%.

Table 1. Distribution of farm sizes (hectares cultivated)

	Cultivated Land Category (% of households)			
Survey Year	**0<X<2**	**2<=X<5**	**5<=X<10**	**X>10**
2012 (full sample)	72.4	23.6	3.3	0.6
2015	61.1	31.8	5.9	1.1

Note: We included the full sample from 2012, including the households that had attrition (i.e., were not included in the 2015 sample)

The second source of data is the annual Zambia CFS. These surveys are conducted in late March - April when the crops have reached physiological maturity. Thus, they reported values are what farmers expected to harvest. A total of six waves is used for this study (i.e. from the 2004/05 agricultural season).

Similar to the RALS data, the CFS data is nationally and district representative of smallholder farm households with a probability proportional to size sample scheme, but unlike the RALS, it is not a panel dataset. Depending on the year, the sample size ranges from 8018 to 13,515, from which we calculate annual district estimates (a total of 72 districts) of the value of crop output per hectare cultivated.[8]

Our lagged productivity variables are calculated by: (a) calculating the gross value of harvested field crops by household using constant per-kilogram 2016 prices; (b) summarizing the median, mean, and median for households cultivating less than 2 hectares and greater than 2 hectares, (c) calculating respective summary measures per district of hectares cultivated by household, (d) dividing each district summary of gross value by the respective district summary of hectares cultivated, giving us a weighted variable representing the value of crops harvest per hectare cultivated.

Finally, we calculate the rainfall measures using the Tropical Applications of Meteorology using Satellite Data and Ground-based Observations (TAMSAT) (Tarnavsky et al. 2014; Maidment et al. 2014; Maidment et al. 2017), which has an approximately 16 square kilometer spatial resolution. We use ArcGIS Model Builder to derive raster (cell based) representations of each measure and to match to RALS household GPS locations.

In table 2 we display the summary statistics, by RALS year, for the dependent variables, all the lags for median district productivity, and the control variables. One thing to note is that the mean adult non-farm income earners are a low share of the mean number of adults in the household, only about 11%.

Table 2. Summary statistics by year: mean and 25th, 50th, and 75th percentiles

Year	2012				2015			
Summary measure	**Mean**	**25pct**	**50pct**	**75pct**	**Mean**	**25pct**	**50pct**	**75pct**
Non-farm income earning "prime age|" Adult household members (No. age 15-59)	0.628	0	1	1	0.716	0	1	1
Non-farm income earning adult equivalents	0.568	0	0.74	1	0.640	0	0.82	1
Key explanatory variables								
Median district crop output (2016 ZMW)/hectare – Lag 1	3,018	2,391	2,875	3,471	3,928	2,228	3,164	5,500
(…) – Lag 2	3,824	2,387	3,483	5,454	3,721	2,251	2,828	5,207
(…) – Lag 3	3,464	1,121	3,151	5,325	3,939	2,685	3,388	4,665
(…) – Lag 4	2,152	1,270	1,801	2,828	3,018	2,391	2,875	3,471
(…) – Lag 5	3,346	1,493	1,747	5,000	3,824	2,387	3,483	5,454
(…) – Lag 6	3,975	831	1,732	6,044	3,464	1,121	3,151	5,325
Control variables								
"Prime age|" Adult household members (No age 15-59)	5.6	4	5	7	6.0	4	6	8
Household adult equivalents	4.5	2.9	4.3	5.8	4.8	3.2	4.6	6.1
Average adult education	5.7	4	6.0	7.5	5.8	4	6	7.5
Max hh education	7.4	6	7	9	7.8	6	8	9
Female headed household (=1)	0.2	0.0	0.0	0.0	0.3	0.0	0.0	1.0
Total land holding size (ha)	2.9	0.9	1.8	3.2	4.1	1.0	2.1	4.3
Tropical livestock units (TLUs)	2.3	0.0	0.0	2.0	2.5	0.0	0.0	2.2
Value of farm equipment (2016 ZMW)	13,090	835	1,943	5,105	15,868	943	2,279	7,269
Growing season (GS) rainfall (mm)	794	722	793	859	849	782	827	912
16-year average of prior GS rainfall (mm)	798	744	813	850	810	753	821	869
GS number of rainfall stress periods (SP)	0.7	0.0	1.0	1.0	1.4	0.0	1.0	2.0
16-year average of prior GS rainfall SP	1.0	0.6	1.1	1.4	0.8	0.4	0.9	1.2
CV of rainfall over previous 16 GS	13.0	10.1	12.7	14.9	11.0	8.7	10.8	12.5

Year	2012				2015			
Summary measure	**Mean**	**25pct**	**50pct**	**75pct**	**Mean**	**25pct**	**50pct**	**75pct**
Number of cell phones/AE	0.2	0.1	0.2	0.2	0.2	0.2	0.2	0.3
Distance to nearest district town (km)	42.1	18.0	35.0	60.0	40.1	17.0	34.0	55.0
Distance to nearest paved road (km)	32.2	5.0	19.0	45.0	29.2	4.0	15.0	42.0
Distance to nearest feeder road (km)	2.0	0.0	0.0	1.0	2.0	0.0	0.0	1.0
Distance to nearest market (km)	26.3	5.0	15.0	35.0	24.5	5.0	14.5	35.0
Distance to nearest agro-dealer (km)	32.2	10.0	24.0	45.0	30.8	8.0	21.0	40.0
Asset Median (<50% = 1)	0.53	0	1	1	0.48	0	0	1

Notes: N = 14,464;

5. RESULTS

In this section, we present the results of our analysis estimating the effect of lagged crop district productivity on the number of household members and the number of adult equivalents receiving non-farm related income. Table 3 presents the results of the Almon lag estimations, including the key control variables, using the median district productivity measures for each district in each lag year. Table 4 presents the full set of long-run results, including both from the Almon lag specification and the moving average specification, and including median, mean, and the median for households cultivating less than 2 hectares, and greater than 2 hectares. Several key results stand out.

First, in the Table 4 long run results, the median and mean results mostly suggest that there is a positive and significant effect of long-run district crop productivity on non-farm labor participation, controlling for the size of the household. The elasticities consistently range from around 0.14 – 0.17, i.e., a doubling of long-run land productivity leads to an increase in non-farm labor participation of 14-17%. The one exception is the mean result in the Almon lagged adult equivalents model, which is not significant at the 10% level.

Second, we find that median productivity among farms that are cultivating less than 2 hectares has a positive and significant impact on non-farm labor – a doubling of productivity leading to a 24-31% increase. On the other hand, the productivity effect of larger farms (2-20 hectares) is either not significant or negative (negative 20-22% for a doubling of productivity). In other words, it is productivity growth on the smallest farms that has the most impact on whether rural farm households are increasing their non-farm sources of income.

Table 3. Elasticities of Multiple Lagged District Median Crop Productivity on the Number of Household Members and Adult Equivalents earning Non-Farm Income (2016 ZMW)

	Total hh members earning non-farm income		Total adult equivalents earning non-farm income	
	Elasticity	*P*-value	Elasticity	*P*-value
District median lag 1	0.002	0.969	-0.003	0.955
District median lag 2	0.106***	0.001	0.102***	0.001
District median lag 3	0.121***	0.001	0.120***	0.001
District median lag 4	0.039*	0.058	0.040**	0.050
District median lag 5	-0.091**	0.049	-0.088*	0.061
District median lag 6	--	--	--	--
District median long-run	**0.174*****	**0.009**	**0.169****	**0.011**
# of HH/AE members	0.328***	0.000	0.467***	0.000
Average adult education	0.129	0.170	0.145	0.122
Max hh education	-0.011	0.906	-0.029	0.763
Female head (=1) (coef.)	-0.049	0.419	-0.087	0.119
Landholding size (ha)	-0.003	0.651	-0.002	0.751
Tropical livestock units	0.009	0.268	0.009	0.331
Value farm equipment	0.002	0.536	0.002	0.519
GS rainfall	-0.184	0.600	-0.122	0.732
16 year mean rainfall	1.137	0.422	0.810	0.563
Rainfall stress periods (SP)	-0.065*	0.072	-0.067*	0.062
16 year mean rainfall SP	0.107	0.547	0.061	0.731
16 year rainfall CV	0.370	0.261	0.344	0.291
Cell phone density	-0.019	0.886	-0.026	0.840
Distance to district town	0.048	0.199	0.046	0.229
Distance to paved road	0.012	0.596	0.020	0.370
Distance to feeder road	0.000	0.937	0.002	0.635
Distance to market	-0.050***	0.002	-0.049***	0.002
Distance to agro-dealer	-0.020	0.348	-0.023	0.282
Asset Median (<50% = 1)	-0.074***	0.000	-0.070***	0.000
Year (2015 = 1) (coef.)	0.367***	0.000	0.312***	0.000

Note: 14,464 observations. We include province dummy variables, province*year interaction effects, and household time-averages of each of the control variables in the models, but they are not reported here. Variables significant at 1%, 5%, and 10%, are denoted by a triple asterisk (***), double asterisk (**), and single asterisk (*), respectively. Elasticities are converted from the initial coefficients that are estimated as levels, except for dummy variables ("Female head" and "Year") which are reported as the initial coefficients.

Table 4. Elasticities of Long-run Crop Productivity (Median, Mean, Median <2 ha and ≥ 2 ha) on the Number of Household Members and Adult Equivalents earning Non-Farm Income, Almon Lag and Moving Average Models (2016 ZMW)

	Total hh members earning non-farm income		Total adult equivalents earning non-farm income	
Almon lag	Elasticity	*P*-value	Elasticity	*P*-value
Median	**0.174*** **	**0.009**	**0.169** **	**0.011**
Mean	0.135*	0.070	0.123	0.100
Median: < 2 ha	0.305***	0.006	0.285***	0.008
Median: ≥ 2 ha	-0.216*	0.073	-0.195*	0.078
Moving average				
Median	0.158**	0.015	0.155**	0.018
Mean	0.153**	0.033	0.135*	0.059
Median: < 2 ha	0.255**	0.020	0.239**	0.029
Median: ≥ 2 ha	-0.169	0.149	-0.158	0.179

Note: 14,464 observations for all median and mean models. 14,280 observations for land size models. We include province dummy variables, province*year interaction effects, and household time-averages of each of the control variables in the models, but they are not reported here. Variables significant at 1%, 5%, and 10%, are denoted by triple asterisk (***), double asterisk (**), and single asterisk (*), respectively. Elasticities are converted from the initial coefficients that are estimated as levels.

5.1. Robustness Checks

We conduct several robustness checks. First, in table A1, we estimate the impact of productivity at the 10th and 90th percentile of productivity in each district, respectively. There is not any evidence that productivity among farms at the 90th percentile in each district affects household non-farm labor. However, there is slight evidence that there is an effect at the 10th percentile in a few of the models, with elasticities around 0.07-0.09. In other words, increases in the productivity of low productivity farms, but not high productivity farms, has a slightly positive impact on non-farm labor.

Second, most of our dataset is based on a two-wave panel, and so we tested each of the models (including at the 10th and 90th percentile) for attrition bias using the Wooldridge (2010) regression approach. As reported in table A2, in all our models we rejected the null hypothesis that there was no attrition bias. In order to investigate whether this attrition bias had a significant impact on the results, we ran the models using an inverse probability weighting (adjusting the weights on the 2015 observations to account for the inverse probability that they would have had attrition). These results are reported in table A3. We found that there was little impact on the overall tenor of the results, both in terms of the magnitude of impact and the statistical significance level. This gives us confidence in the robustness of the initial results, even though there is attrition bias.

Third, we run a similar set of models as reported in table 4 and A1 but exclude all the control variables except those that are rainfall related. This is to test for the possibility that some of the control variables are impacted by the lagged productivity variables, which would result in underestimating their impact on non-farm labor. The results are reported in table A4 and show that the long-run impacts of productivity are mostly lower in magnitude and with less statistical significance. The district-mean results, the median results for farms greater than 2 hectares, and the 10^{th} percentile results that were significant are no longer significant at the 10% level. However, the results for district-median productivity and median for farms less than 2 hectares are still significant. This suggests that any impact of lagged productivity on the control variables does not cause us to under-estimate their impact on non-farm labor.

Fourth, in table A5 we run the land size models with a stricter criterion. In the original models, there are up to 11 districts in any given year that had <10 observations for farms of >=2 hectares cultivated. As a robustness test, we re-ran these models excluding all districts with <10 observations in any of lagged years, a criterion which affected a total of 12 districts and 1,328 observations. We find that the Almon lag results for median land size greater than or equal to 2 hectares is no longer significant, but all results for under 2 hectares still are. This suggests that the initial Almon lag results suggesting that larger farms might have a negative impact on non-farm labor are not very robust and should be interpreted with caution.

6. CONCLUSIONS AND POLICY RECOMMENDATIONS

This is one of the first micro-level empirical studies in sub-Saharan Africa to test the Johnston-Mellor structural transformation hypothesis as it pertains to labor factor market linkages. In particular, we use nationally representative household cross-sectional and panel data to analyze the impact of lagged district crop-productivity on the movement of rural farm household members to off-farm work in the context of Zambia.

There is evidence that the lagged average district-level productivity is associated with an increase in the number of adult household members (or adult equivalents) that receive income from a non-farm (and non-farm wage) source. There is also evidence that this relationship is stronger for changes in productivity among smaller farms (<2 hectares) and farms at the lower end of the productivity distribution (the 10th percentile in each district). It is likely that smaller and relatively less productive farms have a higher income elasticity of demand for basic goods and services that can be acquired in the local economy, something that higher productivity leading to surplus can allow them to fulfill. This, in turn, can lead to more economic activity in the rural non-farm economy, providing opportunities for non-farm labor.

This study has a couple of data-related limitations. First, if we had sufficient data, it would have been useful to evaluate crop labor productivity, not just crop land productivity, which would have provided a more direct link between labor productivity to labor movement into the non-farm sector. Second, based on another limitation in the data, we were only able to include a measure of crop productivity, which did not include the production of fresh fruits and vegetables, and livestock.

Overall, this study appears to align with the structural transformation hypothesis, and in particular the dual economy model of Lewis (1954) and Fei and Ranis (1963) that as regions become more agriculturally productive, farm labor is freed up to pursue income-generating opportunities off the farm. This is important for policy because it suggests that investment in regional agricultural productivity, especially among the smallest farms, can provide opportunities for non-farm labor, allowing a more diversified and resilient livelihood strategy, a dynamic that is highlighted in the literature (Reardon 1997; Barrett et al. 2001; Haggblade et al. 2010; Dedehouanou et al. 2018). While this appears valid for changes in average district productivity overall and for smaller farms, there is also some evidence, less robust, that policy that supports the growth in regional productivity for the least productive farms (relative to each district) can also be helpful in increasing non-farm labor opportunities and enhancing livelihood resilience.

APPENDICES

Table A1. Elasticities of Long-run Crop Productivity (10th and 90th pct) on the Number of Household Members and Adult Equivalents earning Non-Farm Income, Almon Lag and Moving Average Models (2016 ZMW)

	Total hh members earning non-farm income		Total adult equivalents earning non-farm income	
Almon lag	Elasticity	*P*-value	Elasticity	*P*-value
10th percentile	0.089*	0.055	0.070*	0.096
90th percentile	0.027	0.746	0.014	0.874
Moving average				
10th percentile	0.053	0.247	0.078**	0.023
90th percentile	0.097	0.190	0.028	0.705

Note: 14,464 observations for all models. We include province dummy variables, province*year interaction effects, and household time-averages of each of the control variables in the models, but they are not reported here. Variables significant at 1%, 5%, and 10%, are denoted by triple asterisk (***), double asterisk (**), and single asterisk (*), respectively. Elasticities are converted from the initial coefficients that are estimated as levels.

Table A2. Test of Attrition Bias (2016 ZMW)

	Total hh members earning non-farm income		Total adult equivalents earning non-farm income	
Almon	Coef.	*P*-value	Coef.	*P*-value
Median	-0.099***	0.000	-0.090***	0.000
Mean	-0.099***	0.000	-0.090***	0.000
>2 ha, >=2 ha	-0.104***	0.000	-0.094***	0.000
10^{th} and 90^{th} pct.	-0.099***	0.000	-0.092***	0.000
Moving Average				
Median	-0.099***	0.000	-0.090***	0.000
Mean	-0.098***	0.000	-0.090***	0.000
>2 ha, >=2 ha	-0.100***	0.000	-0.091***	0.000
10^{th} and 90^{th} pct.	-0.099***	0.000	-0.091***	0.000

Note: Coefficients estimate the degree of attrition bias, on a dummy variable that is equal to 1 if the household is represented in the 2^{nd} wave. Variables significant at 1%, 5%, and 10%, are denoted by triple asterisk (***), a double asterisk (**), and single asterisk (*), respectively.

Table A3. Main long-run results with Inverse Probability Weighting (IPW) (2016 ZMW)

	Total hh members earning non-farm income		Total adult equivalents earning non-farm income	
Almon lag	Elasticity	*P*-value	Elasticity	*P*-value
Median	0.182***	0.007	0.175***	0.008
Mean	0.146*	0.049	0.132*	0.075
Median: < 2 ha	0.326***	0.003	0.303**	0.006
Median: ≥ 2 ha	-0.221**	0.058	-0.205*	0.080
10th percentile	0.084*	0.069	0.064*	0.124
90th percentile	0.027	0.753	0.022	0.813
Moving average				
Median	0.164**	0.012	0.159**	0.015
Mean	0.158**	0.029	0.139*	0.052
Median: < 2 ha	0.267**	0.015	0.245**	0.025
Median: ≥ 2 ha	-0.176	0.137	-0.159	0.178
10th percentile	0.055	0.217	0.079**	0.036
90th percentile	0.090	0.228	0.037	0.659

Note: 14,464 observations for all median, mean, and 10^{th} and 90^{th} percentile models. 14,280 observations for land size models. We include province dummy variables, province*year interaction effects, and household time-averages of each of the control variables in the models, but they are not reported here. Variables significant at 1%, 5%, and 10%, are denoted by triple asterisk (***), double asterisk (**), and single asterisk (*), respectively. Elasticities are converted from the initial coefficients that are estimated as levels.

Table A4. **Elasticities of Long-run Crop Productivity (Median, Mean, 10th and 90th pct, <2 ha and ≥ 2 ha) on the Number of Household Members and Adult Equivalents earning Non-Farm Income, Almon Lag and Moving Average Models – Rainfall controls only (2016 ZMW)**

	Total hh members earning non-farm income		Total adult equivalents earning non-farm income	
Almon lag	Elasticity	*P*-value	Elasticity	*P*-value
Median	0.137*	0.059	0.134*	0.066
Mean	0.100	0.237	0.087	0.289
Median: < 2 ha	0.240**	0.032	0.218*	0.051
Median: ≥ 2 ha	-0.152	0.207	-0.133	0.272
10th percentile	0.032	0.462	0.045	0.326
90th percentile	0.049	0.617	0.025	0.805
Moving average				
Median	0.139**	0.048	0.137*	0.053
Mean	0.119	0.126	0.101	0.193
Median: < 2 ha	0.212*	0.055	0.196*	0.078
Median: ≥ 2 ha	-0.125	0.297	-0.110	0.362
10th percentile	0.052	0.192	0.056	0.132
90th percentile	0.059	0.505	0.028	0.733

Note: 14,464 observations for all median, mean, and 10th and 90th percentile models. 14,280 observations for land size models. We include province dummy variables, province*year interaction effects, and household time-averages of each of the control variables in the models, but they are not reported here. Variables significant at 1%, 5%, and 10%, are denoted by triple asterisk (***), double asterisk (**), and single asterisk (*), respectively. Elasticities are converted from the initial coefficients that are estimated as levels.

Table A5. Elasticities of Long-run Crop Median Productivity on the Number of Household Members and Adult Equivalents earning Non-Farm Income, Almon Lag and Moving Average Models – Differentiated by Land Size – Stricter Criteria (2016 ZMW)

	Total hh members earning non-farm income		Total adult equivalents earning non-farm income	
Almon lag	Elasticity	*P*-value	Elasticity	P-value
Median: < 2 ha	0.225*	0.065	0.204*	0.096
Median: ≥ 2 ha	-0.184	0.114	-0.168	0.150
Moving average				
Median: < 2 ha	0.246**	0.015	0.231**	0.024
Median: ≥ 2 ha	-0.165	0.135	-0.155	0.170

Notes: We drop any districts that have <10 observations of households cultivating >=2 hectares for any of the lag years (12 out of 72 districts, totaling 1,328 observations). The final models contain 13,136 observations. We include province dummy variables, province*year interaction effects, and household time-averages of each of the control variables in the models, but they are not reported here. Variables significant at 1%, 5%, and 10%, are denoted by triple asterisk (***), double asterisk (**), and single asterisk (*), respectively. Elasticities are converted from the initial coefficients that are estimated as levels.

REFERENCES

Almon, S. 1965. The distributed lag between capital appropriations and expenditures. *Econometrica: Journal of the Econometric Society* 178-196.

Barrett, C.B., T. Reardon, and P. Webb. 2001. Nonfarm Income Diversification and Household Livelihood Strategies in Rural Africa: Concepts, Dynamics, and Policy Implications. *Food Policy* 26 (4): 315–331.

Barrett, C.B., M.R. Carter, and C.P. Timmer. 2010. A century-long perspective on agricultural development. American Journal of Agricultural Economics 92 (2): 447-468.

Block, S., and Timmer, C.P. 1994. Agriculture and economic growth: Conceptual Issues and the Kenyan Experience. A report submitted to the US Agency for International Development (USAID).

Chamberlain, G. 1984. Panel Data. Handbook of Econometrics 2: 1247–1318.

Collier, P., and S. Dercon. 2014. African Agriculture in 50 Years: Smallholders in a Rapidly Changing World? *World Development* 63: 92–101.

Christiaensen, L., and Todo, Y. 2013. Poverty reduction during the rural-urban transformation-the role of the missing middle. The World Bank.

CSO, 2012. 2012 Rural Agricultural Livelihoods Survey Instruction Manual for Listing, Sample Selection, and Largest Maize Field Data Collection. Lusaka, Zambia: CSO.

CSO, MoA, and IAPRI, 2012. 2012 Rural Agricultural Livelihoods Survey Interviewer's Instruction Manual. Lusaka, Zambia: IAPRI.

Dedehouanou, S.F.A., A. Araar, A. Ousseini, A.L. Harouna, and M. Jabir. 2018. Spillovers from off-farm self-employment opportunities in rural Niger. World Development 105: 428-442.

Dennis, B.N., and T.B. İşcan. 2011. Agricultural distortions, structural change, and economic growth: A cross-country analysis. American Journal of Agricultural Economics 93 (3): 885-905.

Dercon, S. and D. Gollin, 2014. Agriculture in African development: theories and strategies. Annu. Rev. Resour. Econ. 6 (1): 471-492.

Diao, X., E.J.Z Robinson, S. Kolavalli, and V. Alpuerto. (2012) The role of agriculture in African economic development: What do we disagree on? In: Jacquet, P., Pachauri, R.K. and Tubiana, L. (eds.) Development, the environment and food: Towards agricultural change? Planet for life. Teri Press, New Dehli. ISBN 9788179934432

Djoumessi, Y.F. Kamdem, C.B. and Nembot, L.N. 2019. Moving off Agrarian Societies: Agricultural Productivity to Facilitate Economic Transformations and Non‐agricultural Employment Growth in Sub‐Saharan Africa. Journal of International Development.

Ellis, F., 2005. Small farms, livelihood diversification, and rural-urban transitions: Strategic issues in Sub-Saharan Africa. The future of small farms 135.

Fei, J.C., and Ranis, G. 1963. Innovation, capital accumulation, and economic development. The American economic review, pp.283-313.

Gujarati, D.N. 2003. Basic econometrics. New York, NY: McGraw-Hill.

Haggblade, S., P. Hazell, and T. Reardon. 2007. Transforming the Rural Nonfarm Economy: Opportunities and Threats in the Developing World. Intl Food Policy Res Inst.

Haggblade, S., P. Hazell, and T. Reardon. 2010. The Rural Off-farm Economy: Prospects for Growth and Poverty Reduction. World Development, 38 (10): 1429–41.

Hazell, P.B. 2013. Options for African agriculture in an era of high food and energy prices. Agricultural Economics, 44 (s1): 19-27.

Hirschman, A.O. 1958. The strategy of economic development (No. 04; HD82, H5.).

Hirschman, A.O. 1992. Linkages in economic development. Rival Views of Market Society and Other Recent Essays 56-76.
IFAD. 2019. Creating Opportunities for Rural Youth: 2019 Rural Development Report, International Fund for Agricultural Development, Rome.
Jacobs, J. 2016. The economy of cities. Vintage.
Johnston, B.F., and J.W. Mellor. 1961. The Role of Agriculture in Economic Development. The American Economic Review 51 (4): 566–93.
Lewis, W.A. 1954. "Economic Development with Unlimited Supplies of Labour." The Manchester School 22 (2): 139–91.
Maidment, R.I., D. Grimes, E. Black, E. Tarnavsky, M. Young, H. Greatrex, R.P. Allan et al. 2017. A new, long-term daily satellite-based rainfall dataset for operational monitoring in Africa Nature Scientific Data 4: 170063 DOI:10.1038/sdata.2017.63.
Maidment, R.I., D. Grimes, R.P. Allan, E. Tarnavsky, M. Stringer, T. Hewison, R. Roebeling, and E. Black. 2014. The 30 Year TAMSAT African Rainfall Climatology and Time Series (TARCAT) Data Set. Journal of Geophysical Research: Atmospheres 119 (18): 610–619.
McKinnish, T. 2008. Panel data models and transitory fluctuations in the explanatory variable. In Modelling and evaluating treatment effects in econometrics, pp.335-358. Emerald Group Publishing Limited.
Mellor, J.W., and J. Lele. 1973. Growth linkages of the new food grain technologies. Indian Journal of Agricultural Economics 18 (1): 35-55
Mellor, J.W. 1976. The new economics of growth. Cornell University Press
Mundlak, Y. 1978. On the Pooling of Time Series and Cross Section Data. Econometrica: Journal of the Econometric Society 69–85.
Pindyck, R., and D. Rubinfeld. 1997. Econometric models and economic forecasts. New York, NY: McGraw-Hill.
Ranis, G., and J.C. Fei. 1961. A theory of economic development. American Economic Review 51 (4): 533-565.
Reardon, T. 1997. Using Evidence of Household Income Diversification to Inform Study of the Rural Nonfarm Labor Market in Africa. World Development 25 (5): 735–47.
Takeshima, H., Amare, M. and Mavrotas, G. 2018. The Role of Agricultural Productivity in Non-farm Activities in Nigeria.
Tarnavsky, E., D. Grimes, R. Maidment, E. Black, R.P. Allan, M. Stringer, R. Chadwick, and F. Kayitakire. 2014. Extension of the TAMSAT Satellite-Based Rainfall Monitoring over Africa and from 1983 to Present. Journal of Applied Meteorology and Climatology 53 (12): 2805–2822.
Tsakok, I., B. Gardner. 2007. Agriculture in economic development: Primary engine of growth or chicken and egg? American Journal of Agricultural Economics 89 (5): 1145-51.
Tschirley, D.L., J. Snyder, M. Dolislager, T. Reardon, S. Haggblade, J. Goeb, L. Traub, F. Ejobi, and F. Meyer. 2015. Africa's unfolding diet transformation: implications for agrifood system employment. Journal of Agribusiness in Developing and Emerging Economies 5 (2): 102-136.
Wooldridge, J.M. 2010. Econometric analysis of cross section and panel data. MIT press.
Yeboah, K.Y., and T.S. Jayne. 2018. Africa's Evolving Employment Trends. Journal of Development Studies 54 (5): 803-832.