I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14864

The Importance of Specification Choices When Analyzing Sectoral Productivity Gaps

Joshua D. Merfeld
Peter Brummund

NOVEMBER 2021

DISCUSSION PAPER SERIES

IZA DP No. 14864

The Importance of Specification Choices When Analyzing Sectoral Productivity Gaps

Joshua D. Merfeld
KDI School of Public Policy and Management and IZA

Peter Brummund
University of Alabama and IZA

NOVEMBER 2021

ABSTRACT

The Importance of Specification Choices When Analyzing Sectoral Productivity Gaps*

A consistent finding in the development literature is that average non-farm labor productivity is higher than average farm labor productivity. These differences in average productivity are sometimes used to promote policies which advance the non-farm sector. In this paper, we analyze the importance of two specification choices when comparing productivity gaps, using detailed household panel data from Malawi. Importantly, we are able to calculate both average revenue products (ARPLs) – similar to most of the sectoral productivity gap literature – as well as marginal revenue products (MRPLs). We show that the choice of productivity measure combined with the choice of production function specification can lead to different sectoral productivity rankings. MRPLs from translog production functions suggest the household farm sector is more productive than the household non-farm sector, while MRPLs from a Cobb-Douglas and ARPLs from both a translog and a Cobb-Douglas find the opposite ranking.

JEL Classification: J24, J43, O13, Q12, R23

Keywords: labor productivity, agriculture, non-farm production

Corresponding author:
Joshua D. Merfeld
KDI School of Public Policy and Management
263 Namsejong-ro
Sejong-si
South Korea
E-mail: merfeld@kdis.ac.kr

* We especially thank two anonymous referees for comments that vastly improved the quality of this manuscript. We also thank Marc Bellemare for comments on a previous version of this paper.

1 Introduction

A large and important debate in development is whether investments in the non-farm sector will be sufficient to pull labor out of agriculture, or whether development need to begin with the agricultural sector. Much of this debate is informed by large differences in average labor productivity across sectors within the same country. This "productivity gap" is present in national accounts data as well as in microdata (**???**). Importantly, these gaps are larger in developing countries than in developed countries, driven in part by large differences in average farm productivity across countries (**?**). These gaps remain, even after taking into account differences in hours worked and human capital across sectors (**??**). However, more recent evidence, using individual fixed effects and identifying gaps based on individuals changing sectors, finds evidence of smaller sectoral productivity gaps (**?**).

While some argue that factor misallocation across firms and sectors may be an important driver of productivity differences (**?????**), differences in average products of labor across firms or sectors do not necessarily imply a misallocation of labor within or across sectors. Indeed, economic theory makes no predictions regarding average products of labor, unless we place strict assumptions on the production functions. Moreover, the importance of using marginal instead of average products is related to the production technology employed by producers. For example, if production follows a Cobb-Douglas process and the labor shares in farm and non-farm production are similar, then the ranking of average products will be the same as the ranking of marginal products. The ranking of average products need not imply anything about the ranking of marginal products if the production technology is not Cobb-Douglas.

In this paper, we contribute to the sectoral productivity gap literature by testing the importance of empirical specification choices in Malawi. Using three waves of household data, we first calculate average revenue products of labor (ARPLs) using similar methods to previous papers. Consistent with the literature, we show that ARPLs in the non-farm sector are much higher than ARPLs in the farm sector. These "productivity gaps" are of a similar magnitude to previous raw gaps calculated using microdata (**??**). We then estimate production functions using the detailed household data. We estimate both Cobb-Douglas and translog production function specifications. We use these production function estimates to calculate marginal revenue products of labor (MRPLs) in both sectors and then compare the resulting estimates with ARPLs.

Consistent with theory, MRPLs calculated from Cobb-Douglas production functions follow similar patterns to the ARPLs: non-farm production appears to be more productive than farm production for the households in the sample. Importantly, however, the translog production function allows us to test whether the Cobb-Douglas specification is flexible enough, since the Cobb-Douglas is a special case of the more generalized translog production function. This test firmly rejects the Cobb-Douglas specification, especially in the non-farm sector. These results suggest that ARPLs and Cobb-Douglas MRPLs may be incorrect, at least in household-level non-farm production.

We come to qualitatively different conclusions when we calculate MRPLs with the translog specification results. Rather than the non-farm sector being more productive, the farm sector is actually more productive. While we are unable to reject equality of MRPLs for farm and non-farm production with the translog results, farm MRPL is almost 30 percent higher than non-farm MRPL. Moreover, the non-farm MRPL drops by more than half with the translog specification relative to the Cobb-Douglas specification. The choice of production function and productivity measure is thus consequential when it comes to estimating sectoral productivity gaps, at least in rural Malawi.

We view our contribution as primarily methodological. Although it is well known that allocative efficiency is related to marginal – not average – productivity, data is rarely sufficient to allow for an empirical test of the importance of these choices. Even in household survey data, few datasets have the required detail, especially in non-farm self-employment, to explicitly test choices of production function specifications and productivity measures. We do not view these results in any way as indicative of economy-wide sectoral productivity gaps, as we are focusing only on household production; large, commercial firms and plantations are absent from our data. Nonetheless, we believe these results are a useful contribution, especially since they speak to production function and productivity estimation in other research areas, not just the sectoral productivity gap literature. Other literature in this general spirit includes **?**, who examine specification choices related to climate change and crop yields; **?**, who look at the importance of data choices related to nightlights; **?**, who discusses specifications in spatial regressions; **?**, who compare specification choices of production in the global agricultural and food system; **?**, who discuss how modeling choices can lead to research disagreement with respect to agriculture and climate change; and **?** who use wages as a proxy for marginal products.

Nonetheless, we do not deny that this paper is related to the sectoral productivity gap literature. While this literature started decades ago – **?**, for example, made this point in his seminal paper – a more recent literature documents continued raw productivity gaps and low productivity in agriculture, especially in developing countries (**?**). Large gaps are also found across urban and rural areas (**???**), which is highly correlated with differences in sectoral productivity. While some of these differences could be driven by differences in human capital (**??**), even gaps adjusted for these differences with microdata suggest differences in sectoral productivity (**??**).

We focus on a different explanation: the choice of productivity measure. Previous literature has noted the shortcomings associated with average products. **?**, for example, use wages to consider productivity differences. However, the wage is not an appropriate measure of marginal productivity when markets are incomplete – which previous literature finds to be the case for Malawi (**?**) as well as in developing countries more generally, e.g. **?** – or when households make labor allocation decisions under risk (**?**). However, one advantage of wages is that they also include employment in larger firms and outside of the household more generally, employment which we do not capture in this paper.

Importantly, while our results differ from some of the previous literature, they do not necessarily imply that sectoral gaps do not exist in Malawi. A key caveat to the results in this paper is that they do not include the largest firms, in both sectors, as we use household survey data. Insofar as larger non-farm firms are more productive than the small household firms in our data, our results are not necessarily generalizable to the economy more broadly. Nonetheless, given the focus on household sectoral choice and poverty, we believe the results here are policy relevant. At the very least, incremental movements of households into the non-farm sector represented here are unlikely to lead to wholesale increases in productivity and, by extension, income or consumption. Nonetheless, it is possible that switching into *different* non-farm jobs, not represented by the households here, could be welfare enhancing.

In the next section, we briefly describe the methods we use to estimate production functions for our sample of Malawian households. We then explore the relationship between productivity and labor allocation. We conclude with some thoughts on the use of average vs. marginal products for cross-country productivity gaps, as well as the choice of production function specification.

2 Methods and Data

2.1 Methods

To calculate marginal revenue products, we estimate both Cobb-Douglas and translog production functions. We choose these two productions functions for several reasons. First, the Cobb-Douglas is arguably the most commonly used production function specification in economics, and the translog is also common. Second, the translog nests the Cobb-Douglas within it, allowing us to test which specification is a better fit for our data. Importantly, we are not able to say whether the translog is the correct production function, per se, but we are able to say whether it is a better choice than the Cobb-Douglas. Finally, both specifications are empirically tractable. Another option, for example, would be a constant elasticity of substitution (CES) production function. However, with so many household fixed effects, convergence is problematic.[1] Moreover, the translog specification is a linear approximation of the CES, so it is related.

We estimate translog production functions of the form:

$$\ln R_{iht} = \alpha_h + \sum_j \beta_j \ln \gamma_{jiht} + \sum_j \sum_k \beta_{jk} \ln \gamma_{jiht} \ln \gamma_{kiht} + \delta C_{iht} + D_{dt} + \eta_m + \varepsilon_{iht}, \qquad (1)$$

where the key variables are revenue (R_{iht}) and productive inputs (γ_{jiht} and γ_{kiht}, where j and k both index different productive inputs, allowing for squared terms and interactions between different inputs), on enterprise i in household h in time t. For farm production, we define labor, land, and fertilizer as productive inputs. For non-farm production we include just labor and total non-labor costs as productive inputs. For both farm and non-farm production, we use revenue as our output measure to enable comparisons across different

[1] We encountered issues with convergence with even simpler specifications than the final specifications we use in this paper.

types of activities with different types of outputs. Due to possible price differences across time and space – which may bias production function estimates with costs as dependent variables (**?**) – we control for district/wave fixed effects (D_{dt}) and month of interview (η_m) fixed effects. Finally, C_{iht} is a vector of controls that may affect revenue and which differ depending on the sector of the enterprise. We estimate these production functions separately for both the farm and non-farm sectors.

The Cobb-Douglas production function specifications are similar:

$$\ln R_{iht} = \alpha_h + \sum_j \beta_j \ln \gamma_{jiht} + \delta C_{iht} + D_{dt} + \eta_m + \varepsilon_{iht}, \tag{2}$$

where the only difference is that we do not include the higher-order terms involving the productive inputs. Note that this Cobb-Douglas production function is nested within the translog production function in Equation 1. A simple test for the appropriateness of a Cobb-Douglas specification – at least relative to a translog specification – is an F-test of joint significance for all higher-order terms. This motivates our decision to test the Cobb-Douglas – one of the most common production function specifications – against the translog specification. Thus, while we cannot definitively say that production in these households follows a translog production function, we can say whether a translog specification is a better option than the Cobb-Douglas specification for each production activity.

We rely on household fixed effects for identification. In a related paper, using a similar identification strategy, we argue that endogeneity of input choices is unlikely to explain the differences we observe between farm and non-farm productivity **?**[2]. Importantly, any endogeneity affecting our MRPL estimates can also affect the ARPL estimates.[3] As such, we see this paper as contributing to the sectoral productivity gap literature – even though

[2]This includes the concerns raised by Gollin and Udry (**?**).

[3]For example, any unobserved production factor will not be properly subtracted from total revenue when creating a labor "value added" estimate, incorrectly attributing that additional productivity to only labor.

we are not attempting to make a general comment on the size of that gap here – where a comparison of average products is common. Specifically, average products are often calculated by first subtracting all costs from total production. However, if we do not observe all relevant costs or inputs, then the estimated average product of labor may be overestimated and misleading.

For average product estimates, we follow **??** and construct a "value added" for labor by subtracting all explicit costs from total revenue. However, we do not subtract hired costs, since we pool family and hired labor together in the production function estimation.

2.2 Data

We use data from three waves of the Malawian Integrated Household Survey (IHS): the 2010, 2013, and 2016 waves. The IHS is collected by the Malawian National Statistical Office, with support provided through the Living Standards Measurement Study program at the World Bank. There is a small panel component that is interviewed in all three waves, but the majority of observations come from large cross-section components of the 2010 and 2016 waves. The data is nationally representative, but is obviously lacking information on large non-farm firms, which are arguably the most productive non-farm entities. The data also does not capture large commercial farms. As such, our overall results are not directly comparable with much of the sectoral productivity gap literature. Nonetheless, our results with average products are in line with recent evidence from microdata (**??**).

In order to compute revenue in agriculture, we construct geographic prices for each crop. We restrict attention to plots grown with only crops for which we have enough observations to construct these prices: maize, tobacco, groundnut, rice, sweet potato, potato, beans, soya, pigeon peas, cotton, sunflower, pumpkin leaves, and tomato.[4] We check ro-

[4]We include only the labor that is allocated to these plots. As such, while we are not using the entirety

bustness to different price requirements. For example, we construct aggregate prices with no restriction on the number of observations – that is, taking the first price we observe at the lowest level of geographic aggregation – requiring at least five price observations, and requiring at least ten price observations. These different restrictions have no affect on our conclusions. All results we present in this paper use the second requirement, of at least five price observations.

To give a sense of the geographic aggregations with this requirement, in the first wave, 5,851 plot-crop price observations use an aggregate price at the enumeration area level, which is quite small, similar to a small village in rural areas. 18,011 plot-crop price observations are constructed using district-level prices and 8,882 plot-crop price observations use an aggregate price at the region level. We never assign prices at the country level and drop remaining plots with any missing price observations. If we require just a single price observation, almost 18,000 plot-crop prices are assigned at the EA level, 10,143 at the district level, and 9,431 at the region level.[5]

For non-farm production, households were asked specifically about total revenue and total costs. As such, we do not need to construct aggregate prices for non-farm production. Costs include raw materials, transportation costs, fuel/oil/electricity/water, insurance, and "other."' We deflate all revenue and costs, for both farm and non-farm production, to 2010 MWK.[6]

Across the three waves of data, we observe a total of 7,030 non-farm enterprises and 34,353 agricultural plots, after a few restrictions that we enumerate below. While we use

of agricultural output for all farms, we should not be underestimating labor productivity since we are also purposefully underestimating labor in the exact same way. We likewise do not include livestock revenue given the difficulties related to properly assigning productivity measures. For example, if a household grazes some goats but does not explicitly consume meat or milk, it is not clear how to assign revenue even if the goats act as a buffer stock and/or they will provide explicit consumption in the future.

[5]Requiring just a single price observation allows for more non-missing price observations, but at the cost of likely introducing substantial error in the assignment of prices.

[6]The exchange rate in 2010 was approximately 120-125 MWK per USD.

this entire sample to estimate average products and marginal products, not all of these households contribute to identification due to the use of household fixed effects. In terms of identification, generally only households that have at least two observations in a given sector will contribute to identification. There are 2,036 non-farm enterprises across 873 different households and 25,367 agricultural plots across 8,801 different households that fit this criterion. The difference in numbers is because many households operate more than one plot, whereas many households operate just a single non-farm enterprise. Since the bulk of our sample is from the cross-section waves, this leads to a much larger agricultural sample.

For both farm and non-farm production, we pool hired and household labor because of the low levels of hired labor in our sample, which is confirmed in Table 1. Specifically, households report hiring any labor for around 17.7 percent of plots. However, this hired labor accounts for only 3.5 percent of total labor days across all agricultural plots in our sample. While non-farm enterprises hire less often – just 10 percent reported hiring anyone – those that do hire employ slightly more hired labor; hired labor makes up around 12 percent of all labor in non-farm enterprises in our data. We include both types of labor in estimation, we just pool them together since we do not have enough hired labor to separate out into its own category.[7]

For average products, we calculate revenue net of non-labor costs, similar to **?** and **?**. We implicitly treat hired labor the same as household labor in the average product calculations, for a more direct comparison with the marginal products.

Focusing on non-farm production for simplicity, we can calculate MRPL from the translog

[7] **?**, using different data, report that agricultural wages make up around 10-15 percent of rural household income. The discrepancy between those numbers and the numbers in the LSMS data employed here is likely explained by a difference in sample. Specifically, our data do not include large commercial farms, which are likely to hire large amounts of labor. As such, the households in our sample are more likely to sell their labor on the market than to buy labor on the market.

specification as:

$$MRPL^{tl}_{nf} = (\beta_L + 2\beta_{LL} \log L_{iht} + \beta_{LC} \log C_{iht}) \frac{\hat{R}_{iht}}{L_{iht}}, \tag{3}$$

where β_L is the coefficient on labor, β_{LL} is the coefficient on labor squared, β_{LC} is the coefficient on the interaction between labor and input costs, L_{iht} is total labor allocation in days, and C_{iht} is total input costs. R_{iht} is predicted revenue.[8]

In the Cobb-Douglas case, the marginal product is always directly proportional to the average product, and is calculated as:

$$MRPL^{cd}_{nf} = (\beta_L) \frac{\hat{R}_{iht}}{L_{iht}}. \tag{4}$$

Finally, we are interested in comparing productivity across activities at the household level. However, estimation is at the non-farm enterprise/plot level. As such, we need to aggregate enterprise-level results to the household. Here, we take a simple median across plots/enterprises. An alternative option, weighting MRPL across plots/enterprises by labor allocation, does not affect conclusions (**?**). To construct standard errors for MRPLs and differences, we bootstrap the entire process 1,000 times. We set up the bootstrap to draw households, due to the use of household fixed effects and clustered standard errors.

2.3 Summary statistics

Table 1 presents farm and non-farm summary statistics. The first two columns present farm statistics, while the last two columns present non-farm summary statistics. The first and third columns are at the plot/enterprise-wave level and the second and fourth columns are at the household-wave level. The first thing to note is how many households operate

[8]The choice of predicted vs. actual revenue does not affect qualitative conclusion.

multiple plots, even after restricting the sample to the 13 most common crops. In non-farm production, on the other hand, most households operate just a single enterprise in any given wave. This could make aggregation an important consideration, as just discussed, but in practice the method of aggregation of MRPLs to the household level has no effect on qualitative results, as shown in **?**.

The second thing to note is how little hired labor households in our sample employ in both farm and non-farm production. In farm production, at the household-wave level, fewer than five percent of all labor days are hired labor days. In non-farm production, hiring is apparently slightly more prevalent, but even there hired days make up just over 12 percent of total days at the household level. Interestingly, the percentage of households who hired any labor at all is somewhat higher in farm production, at 22.7 percent of households. However, among households who hired any labor for farm production, the median is just 10 days. Among all households, the median is zero.

Figure 1 shows the breakdown of industries in our household data. The most common industry is petty trade and restaurants – making up more than 60 percent of firms – followed by manufacturing. The other industries make up a relatively small proportion of firms in our data.[9] An important caveat to our results is that household enterprises are not a random sample of all non-farm firms. As such, our results do not necessarily represent the larger firms in Malawi. Nonetheless, since much of the discussion around development policy relates to policies affecting individual households and occupational choices, including types of self-employment, we believe the results are relevant.

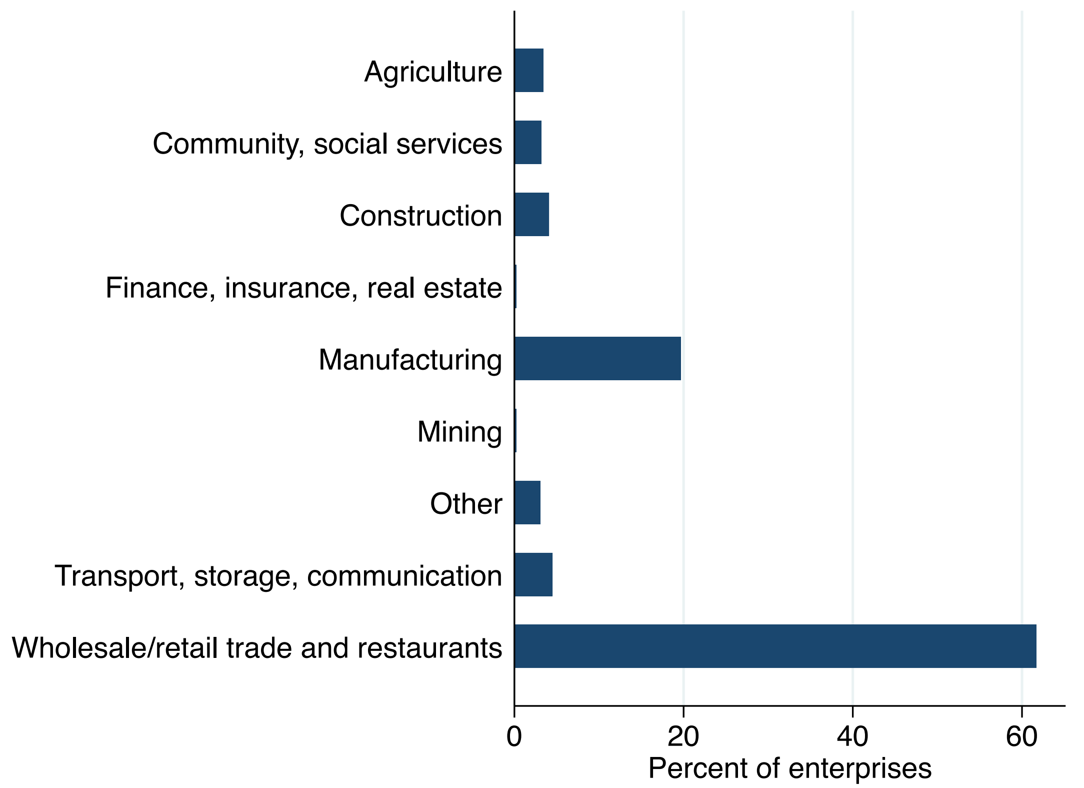
[9]The few non-farm enterprises that list “agriculture” as their industry is somewhat puzzling. We are not exactly sure of the nature of these enterprises, but they are a very small proportion of all enterprises. In the raw 2010 data, for example, just 2 out of 2,809 household non-farm enterprises are listed as “agricultural” enterprises.

Table 1: Summary statistics - Plot/enterprise and household levels

	Farm		Non-Farm	
	(1)	(2)	(3)	(4)
	Plot level	HH-wave level	Enterprise level	HH-wave level
	mean/sd	mean/sd	mean/sd	mean/sd
Revenue (MWK)	30,321	58,561	31,245	37,439
	(58,134)	(124,697)	(70,033)	(89,213)
Total labor (days)	61.9	120	23.9	28.6
	(43.6)	(117)	(19.6)	(28.0)
Family labor (days)	59.1	114	20.9	25.2
	(43.7)	(115)	(15.4)	(22.0)
Hired labor (days)	2.16	4.17	2.95	3.54
	(8.04)	(15.9)	(12.3)	(14.8)
Hired any labor (yes=1)	0.177	0.227	0.097	0.106
	(0.382)	(0.419)	(0.296)	(0.308)
Total wages (MWK)	1,800	3,476	866	1,037
	(55,524)	(79,000)	(6,414)	(7,759)
Acres	1.09	2.11		
	(10.4)	(22.0)		
Fertilizer (kg)	11.3	21.7		
	(26.3)	(48.3)		
Total non-labor costs			22,094	26,474
			(78,029)	(95,867)
Observations	34,353	17,787	7,030	5,867

Standard deviations are in parentheses. Farm statistics are for the previous season, while non-farm statistics are for the previous month of operation.

Figure 1: Household enterprise industries



3 Results

We start with estimates of average revenue products in Table 2. ARPL in non-farm production is higher than in farm production, whether we use the mean or the median. For the mean, non-farm productivity is approximately 3.5 times higher than farm productivity, while at the median it is about 2.7 times higher. Compared to previous estimates, these numbers are similar. **?** estimate a mean raw productivity gap – unadjusted for differences in time worked in a sector – of around 3.5 across all countries for which they have data. Using data from Malawi in 2005, those authors estimate that primarily agricultural-sector workers spend approximately 26.4 hours in the sector, while primarily non-agricultural workers spend approximately 38.2 hours in the sector. Using these numbers, a naive adjusted mean gap in our data would be almost 2.5. Using the 2010 wave of the Malawi data, **?** estimates that differences in hours worked may decrease the gap to as little as 1.4. As a robustness check below, we also estimate these gaps only for households that are engaged in both sectors simultaneously.

Table 2: Mean and Median of Average Revenue Product of Labor

	(1) Mean	(2) Median
Farm	584	242
Non-Farm	2,077	655

All statistics are calculated as revenue in MWK over number of days worked. Actual revenue uses reported revenue (non-farm) or constructed revenue using aggregate prices and reported harvest (agriculture). For both cases, costs are subtracted from total revenue, except for wages. Since we treat hired and family labor equally in the MRPL calculations, we also do so here.

To calculate marginal products, we first need to estimate production functions. Table 3 presents these estimates. Column one presents farm results using a Cobb-Douglas specification, while column two presents farm results from a translog specification. Columns three and four present non-farm results from a Cobb-Douglas and translog specification, respectively.

There are several things of note. First, the labor coefficients in the Cobb-Douglas specification are quite similar across the two household sectors, which would lend support to the use of average products in comparing productivity across activities.[10] In fact, we are unable to reject equality of the two coefficients at traditional levels of significance (p=0.109). This may be important, given that average and marginal products are proportional in a Cobb-Douglas specification, a point to which we return below.

Second, the test for a nested Cobb-Douglas rejects a Cobb-Douglas specification for both farm and non-farm production. However, the coefficients suggest differences in the non-farm sector may be more likely to hinge on labor. Both higher-order terms with labor (labor squared and labor times costs) are at least twice as large in magnitude as the largest higher-order term involving labor in the agricultural translog specification. Take, for instance, the interaction between non-farm labor and total costs, which is negative. In Equation 3, part of the MRPL calculation involves subtracting the product of the coefficient on the interaction term and total costs for the household. Evaluated at median non-farm costs (8.51), this product is equal to 0.33, which is approximately equal to the sum of the non-farm labor term (0.255) and the non-farm labor squared term times two (0.033). Costs play a major role in mediating non-farm MRPL in the translog specification, but do not play a similar role in the Cobb-Douglas. Whether this affects results remains to be seen.[11]

Finally, Table A1 in the appendix presents production estimates after first collapsing the

[10]Recall that the MRPL for a Cobb-Douglas is defined as $(\beta_L)\,\frac{\hat{R}_{iht}}{L_{iht}}$, where β_L is the coefficient on (log) labor from the production function. If labor shares are similar, using the average product will yield similar conclusions regarding productivity differences across non-farm and farm production as using the marginal product. However, if the labor shares are markedly different, using average products may still be misleading with the Cobb-Douglas specification.

[11]While the coefficient on the level term for costs is negative, the squared term is positive, meaning the shape implied by the coefficients – ignoring labor for simplicity – is an upwards-opening parabola. The vertex of that parabola if labor were zero is just 2.74; less than eight percent of enterprises in our sample has costs below this number. In other words, for almost all enterprises in our sample, increasing costs is associated with increasing output.

data to the household-wave level (that is, they are household-level production functions), separately for farm and non-farm production. The overall conclusions are similar to the results presented in Table 3, though there are some differences. Interestingly, the non-farm results are actually more precisely estimated at the household level. This makes both of the non-linear terms significant and also makes clear that the Cobb-Douglas may not be appropriate for non-farm production.

We clearly reject a Cobb-Douglas specification for both, but especially for non-farm production. This suggests that average products may not be appropriate for comparing sectoral productivity. We delve further into this possibility in Table 4, where we calculate MRPLs using the production function results from Table 3. The first column presents MRPLs from columns one and three of Table 3, the Cobb-Douglas specifications. Results are consistent with the average products presented in Table 2, especially after taking into account some differences in time worked across sectors. The MRPL results in column one suggest that the non-farm sector is around 1.6 times more productive than the non-farm sector.

Column two presents MRPL estimates using translog production function specifications. The results are markedly different from those in column one. While we cannot reject equality of the MRPL estimates, non-farm MRPL is now *lower* than farm MRPL, by almost 30 percent. Moreover, the difference in column three, which shows the change in MRPL when moving from the Cobb-Douglas to the translog specification, firmly rejects equality of non-farm MRPL across specification choices. Interestingly, farm MRPL is almost identical, with a difference of just 0.1 *percent* across specifications.

We can put these MRPL estimates in context using the 2010 exchange rate from MWK to USD. At the time, around 120-125 MWK were equal to one USD. The MRPL estimates in column one suggest an additional person-day in the farm sector would result in around an

Table 3: Production Function Estimates

	Farm		Non-farm	
	(1)	(2)	(3)	(4)
	C-D	Translog	C-D	Translog
Labor (L)	0.256***	0.175*	0.192**	0.255
	(0.020)	(0.106)	(0.096)	(0.279)
Acres (A)	0.374***	0.471***		
	(0.018)	(0.100)		
Fertilizer (F)	0.083***	0.217***		
	(0.008)	(0.057)		
Costs (C)			0.251***	−0.340***
			(0.052)	(0.110)
$L \times L$		0.010		0.033
		(0.014)		(0.055)
$A \times A$		0.015		
		(0.017)		
$F \times F$		−0.023***		
		(0.008)		
$L \times A$		−0.020		
		(0.019)		
$L \times F$		−0.007		
		(0.010)		
$F \times A$		−0.003		
		(0.011)		
$C \times C$				0.062***
				(0.009)
$L \times C$				−0.039
				(0.030)
F-test C-D (p)		0.059		0.000
Observations	34,314	34,314	7,026	7,026

Standard errors clustered at the household level are in parentheses. Household fixed effects are included in all regressions. Also included are month of interview fixed effects and wave/district fixed effects. In addition, we include crop dummies, plot quality variables, non-farm industry dummies, and a dummy indicating whether the non-farm industry has access to electricity. The F-tests present tests for a nested Cobb-Douglas production function in each translog; the p-value is constructed by testing whether all squared and interaction terms are simultaneously zero. Revenue and non-farm costs are in (March) 2010 MWK.
* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table 4: Median MRPL Estimates

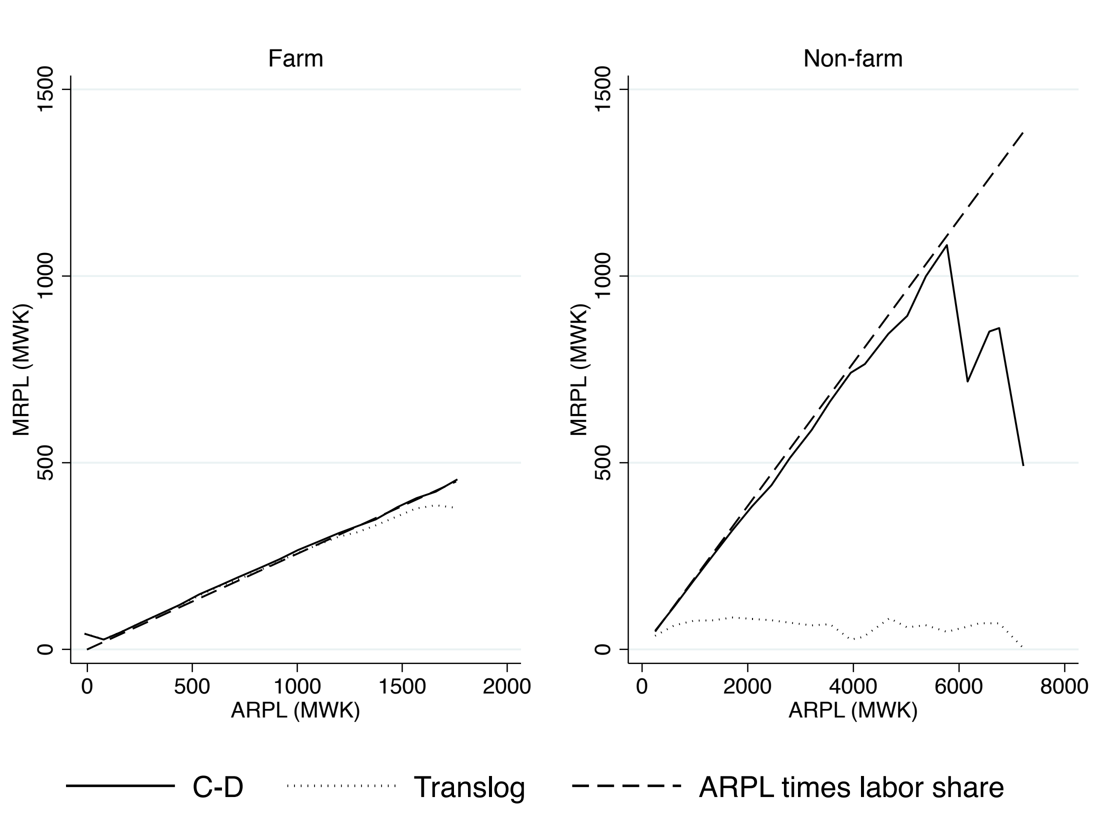
	(1) Cobb-Douglas	(2) Translog	(3) Difference
Farm	72.265***	71.912***	0.352
	(3.915)	(4.340)	(2.312)
Non-farm	114.049***	51.005***	63.044***
	(23.421)	(21.051)	(19.605)
Difference	−41.784*	20.908	
	(23.735)	(21.532)	

The first column presents MRPLs constructed using production function estimates from columns one and three of Table 3. The second column presents MRPLs constructed using production function estimates from columns two and four of Table 3. All standard errors are constructed through bootstrapping the MRPL construction 1,000 times. The bootstrap is set to draw households. MPRL estimates are in 2010 MWK. MRPL difference (row three) is constructed as farm MRPL minus non-farm MRPL. Column three is constructed as Cobb-Douglas minus translog.
* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

additional 0.59 USD, while an additional day in the non-farm sector would return approximately 0.93 USD. This suggests that a household moving from the farm to the non-farm sector would increase its daily return by more than 0.3 USD. To put this number in further context, using the 2010 IHS data, we can calculate average household income per capita in the full sample. Across all households in the 2010 IHS, this comes to an average of 176.13 MWK per person per day. In other words, the MRPL difference is equal to almost one-quarter of total average daily income per person, suggesting considerable potential income is left on the table by not equating MRPL across activities. However, when using the translog specification, we see that households moving from the farm to the non-farm sector might actually lead to a decrease in income, not an increase. This highlights the importance of the specification choices.

To see how these estimates vary with average products, we present these numbers graphically in a median band plot in Figure 2. The left figure is for the farm sector and the right figure is for the non-farm sector. On the x-axis in both figures is a household's average product in a given sector. On the y-axis is MRPL, which we calculate in three ways: 1) using the Cobb-Douglas specification (column one of Table 4); 2) using the translog

Figure 2: Average products vs. marginal products, by sector



The left figure presents farm results and the right figure presents non-farm results. To improve presentation, only the middle 90 percent of MRPLs are included. Both plots are median band plots.

specification (column two of Table 4); and 3) multiply ARPL by the labor share from the Cobb-Douglas specification. We present only the middle 90 percent of the ARPL/MRPL distribution in order to ease presentation.

There are several interesting features in the graph. Perhaps most striking is how closely the Cobb-Douglas MRPLs track average products times the labor share, in both the farm and non-farm sectors. While this is theoretically expected, we calculate average products differently from marginal products and, as such, the relationship shown in Figure 2 suggests the methodology is appropriate. Second, the choice of a Cobb-Douglas or translog specification is apparently completely inconsequential for the farm sector. Across all range of household ARPL, the two MRPL measures are almost identical. Finally, the use of translog MRPL in the non-farm sector has large impacts on qualitative conclusions. Cobb-Douglas and translog MRPL begin diverging almost immediately, and continue to do so for most of the distribution.

As a check on these results, we also re-estimate production functions at the household level. We are principally interested in household-level labor allocation across activities, not plot- or enterprise-level labor allocation, which makes household-level estimation attractive. Moreover, the preceding results require us to aggregate across different plots/enterprises, which is not straightforward. The downside to collapsing to the household before estimating production functions, however, is that we lose quite a few observations and power. We present these results in Table 5. There are several things to note. First, the qualitative conclusions are unchanged; although non-farm MRPL is slightly higher – though not significantly so – when calculated using a Cobb-Douglas specification, farm MRPL is significantly higher than non-farm MRPL – by around four times at the median farm and non-farm MRPL, but by around 2.5 times for the median difference.

Second, when looking at the third column, it is striking how precisely estimated the difference in farm MRPL is across specifications. This is likely due to the fact that translog and Cobb-Douglas specifications are both reasonable choices for farm production, meaning that their MRPL estimates are quite similar, even across repeated bootstrap replications.

The plot/enterprise-level and household-level production functions result in similar conclusions but the magnitudes are somewhat different. It is difficult to know the exact reason for this, but it is worth noting that the household-level production functions make assumptions regarding production technologies across household plots or enterprises that the more disaggregated production functions do not. In addition, we lose some households in the household-level production function approach. For example, the two large cross-section

Table 5: Household-Level Production Function Estimation
Median MRPL Estimates

	(1) Cobb-Douglas	(2) Translog	(3) Difference
Farm	105.470***	93.085****	12.384***
	(6.194)	(6.352)	(3.129)
Non-farm	116.761***	26.358**	90.403***
	(13.680)	(11.668)	(9.991)
Difference	−11.291	66.727***	
	(15.205)	(13.123)	

All columns present results from production functions estimated from data collapsed to the household-wave level. All standard errors are constructed through bootstrapping the MRPL construction 1,000 times. The bootstrap is set to draw households. MPRL estimates are in 2010 MWK. MRPL difference (row three) is constructed as farm MRPL minus non-farm MRPL. Column three is constructed as Cobb-Douglas minus translog.
* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

IHS samples in 2010 and 2016 have households that we observe only once. If these households have two plots and no non-farm enterprises, then they contribute to identification if we estimate production functions at the plot level but not if we estimate production functions are the household level.

We also present results for households in both sectors by estimating household-level production functions. These results are in Table A2 of the appendix. Results are again consistent with the primary results presented above. We are unable to reject MRPL equality using C-D production functions, but we strongly reject equality when using the translog production functions.

One important difference across sectors, emphasized by results in both **?** and **?**, is that human capital differences across sectors may be an important driving force of any productivity differences. While we find differences in the opposite direction, differences in human capital may still be present. Ideally, we would be able to estimate marginal revenue products of labor for individuals within households. We could in theory do this for plots and enterprises for which a single person works. However, while there are some enterprises

Table 6: Median MRPL Estimates, households in both sectors

	(1) Cobb-Douglas	(2) Translog	(3) Difference
Farm	112.354***	115.306***	−2.952
	(13.301)	(14.398)	(5.521)
Non-farm	68.456**	6.959	61.497**
	(30.446)	(17.531)	(24.557)
Difference	43.898	108.347***	
	(30.446)	(22.256)	

Estimation includes only households that are engaged in both the farm and non-farm sector in the same wave. The first column presents MRPLs constructed using production function estimates from columns one and three of Table 3. The second column presents MRPLs constructed using production function estimates from columns two and four of Table 3. All standard errors are constructed through bootstrapping the MRPL construction 1,000 times. The bootstrap is set to draw households. MPRL estimates are in 2010 MWK. MRPL difference (row three) is constructed as farm MRPL minus non-farm MRPL. Column three is constructed as Cobb-Douglas minus translog.
* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

operated by just a single individual, most plots have more than one individual working on them throughout a season.

Given that we are interested in household-level labor allocation, we instead think of this at the household level, and re-restimate production functions and MRPLs for only households that are engaged in both sectors. Since these households are engaged in both sectors simultaneously, human capital differences at the *household* level are minimized. Nonetheless, individual-level differences could, in theory, still explain differences, especially in a “separate spheres" style decision-making process.

We present these results in Table 6. These results are related to, and somewhat replicate, tests for efficiency in **?**.[12] While farm MRPL is higher using both production functions for this subgroup of households, the overall effect of using a translog or Cobb-Douglas specification is similar. Apparently, the choice is again inconsequential for the farm sector, but leads to large differences in estimated MRPL in the non-farm sector.

Finally, an alternative to estimating MRPLs is to use wages to infer the opportunity cost

[12]The results in column two are identical to a similar table in **?**. That paper does not estimate Cobb-Douglas MRPLs, however.

of time. However, wages represent the relevant opportunity cost of time only when markets are complete and there are no frictions. In Malawi, specifically, there is evidence of these types of frictions. **?** and **?** both find evidence of market failures in several Sub-Saharan African countries, including Malawi. **?** also finds specific evidence of labor market frictions leading households to overallocate to own production, especially in the lean season. We can examine whether this is consistent with our results, as well, by comparing estimated MRPLs to reported wages. In our sample, the average wage paid by households – across both farm and non-farm enterprises – to hire labor from the market is around 343 MWK, a number that is decidedly higher than any of the estimated MRPLs.

4 Conclusion

A large body of literature finds that the non-farm sector as a whole is more productive than the farm sector. These findings often lead to the policy conclusion that incentivizing households to move from one sector to the other could lead to significant increases in incomes for both individual households and, as such, the entire economy. However, this literature primarily compares average products of labor across sectors. The results we present here, however, indicate that the use of average products may not be justified in making these comparisons, at least in the case of households in Malawi. This paper emphasizes the importance of two specification choices when comparing labor productivity: the specification of the production function and the choice of using marginal or average products. For households in Malawi, average revenue products of labor are higher in non-farm production than in farm production. However, average products are only defensible if production follows a Cobb-Douglas process and labor shares are equal across sectors; we show that neither is true in our data. Consistent with the average product results, marginal revenue products constructed from Cobb-Douglas production function specifications. We find the opposite when using marginal revenue products based on translog production functions: MRPLs are higher in farm production

Importantly, the difference between the Cobb-Douglas and translog specifications is only apparent for non-farm enterprises. This may be related to the large amount of heterogeneity in the non-farm sector in developing countries (**?**). Some households may operate relatively unproductive enterprises – with lower capital requirements – while others may operate more capital-intensive enterprises. Since the underlying production processes may be quite different for these different types of enterprises, the quadratic terms in the translog specification may better accommodate variation in input shares while the interaction terms may better capture input complementarities and substitution.[13]

This paper uses household microdata and, as such, it does not include many of the most productive firms or farms. In other words, the results presented here are not necessarily at odds with the macro sectoral productivity gap literature. Any potential lack of representativeness of our data does not compromise the internal validity of our findings regarding the importance of specification choices. Nonetheless, our average product estimates of productivity gaps are relatively similar to those calculated using microdata, such as **??**.

In other words, it is not clear that non-farm household enterprises are more productive than household farm enterprises, at least in Malawi.[14] While much of the policy dialogue is focused on cross-section reallocation of labor (from the farm to the non-farm sector), our results suggest that a reallocation towards the farm sector within households may also be capable of improving household incomes. Whether this type of reallocation could lead to a broader process of structural transformation is yet to be determined (**?**).

We find evidence that the non-farm sector is, on average, more productive than the farm sector, but that the farm sector is more productive at the margin, at least for certain house-

[13]We are especially appreciative to an anonymous referee for this insight.
[14]Again, this says nothing about large firms in either sector.

holds. Importantly, at the household level, even differences in marginal revenue products need not imply misallocation. After all, households facing risk can make perfectly rational decisions that are not consistent with MRPL equality (**?**). Nonetheless, there is a vast literature on misallocation, both across and within sectors (e.g. **??????**). As such, we do not argue that differences in MRPL are necessarily signs of misallocation. Rather, we instead argue that it is important to test implicit assumptions regarding production functions and to use theoretically appropriate measures of productivity. While the Cobb-Douglas is prized for its simplicity, the assumption that production follows a Cobb-Douglas relationship is not trivial, as we show here. Moreover, the choice of production function specification can have important implications for empirical results.

It is not clear that labor productivity in household enterprises is more productive than in small farms, at least not in Malawi (results suggest the opposite, given that the use of average revenues is rejected as criterion (labor shares aren't equal) and given that the data underscore the need for using translog specifications, at least when looking at household enterprise data. These findings are consistent with the recent shift in emphasis on economic as opposed to structural transformation in the policy dialogue, with the former referring to the importance of within sectoral transformation towards more efficient farm and non-farm enterprises (and the latter more often used to refer to cross-sectoral movements of production factors) (e.g. Fuglie et al 2020, but also ACET in Ghana). Both will be needed, but a lot remains to be gained from within sector improvements.

Appendix

Table A1: Household-level production functions

	Farm		Non-farm	
	(1)	(2)	(3)	(4)
	C-D	Translog	C-D	Translog
Labor (L)	0.248***	−0.016	0.150**	0.360**
	(0.025)	(0.117)	(0.049)	(0.145)
Acres (A)	0.410***	0.374***		
	(0.033)	(0.133)		
Fertilizer (F)	0.066***	0.336***		
	(0.010)	(0.062)		
Costs (C)			0.263***	−0.291***
			(0.029)	(0.049)
$L \times L$		0.037***		0.005
		(0.013)		(0.028)
$A \times A$		−0.014		
		(0.018)		
$F \times F$		−0.029***		
		(0.008)		
$L \times A$		0.002		
		(0.028)		
$L \times F$		−0.028**		
		(0.011)		
$F \times A$		0.009		
		(0.015)		
$C \times C$				0.056***
				(0.004)
$L \times C$				−0.037***
				(0.012)
F-test C-D (p)		0.556		0.003
Observations	4,782	4,782	1,127	1,127

Standard errors clustered at the household level are in parentheses. The results here are analogous to those presented in Table 3, except we first collapse to the household-wave level before estimation. The F-tests present tests for a nested Cobb-Douglas production function in each translog; the p-value is constructed by testing whether all squared and interaction terms are simultaneously zero. Revenue and non-farm costs are in (March) 2010 MWK.

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

Table A2: Median MRPL Estimates, households in both sectors
Household-level production functions

	(1) Cobb-Douglas	(2) Translog	(3) Difference
Farm	135.985***	101.076****	34.909***
	(11.044)	(10.550)	(6.439)
Non-farm	125.123***	37.080***	88.043***
	(14.381)	(13.660)	(10.441)
Difference	10.862	63.996***	
	(18.499)	(13.123)	

Estimation includes only households that are engaged in both the farm and non-farm sector in the same wave. All columns present results from production functions estimated from data collapsed to the household-wave level. All standard errors are constructed through bootstrapping the MRPL construction 1,000 times. The bootstrap is set to draw households. MPRL estimates are in 2010 MWK. MRPL difference (row three) is constructed as farm MRPL minus non-farm MRPL. Column three is constructed as Cobb-Douglas minus translog.
* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$