



INTERNATIONAL
FOOD POLICY
RESEARCH
INSTITUTE

IFPRI



RESEARCH
PROGRAM ON
Policies,
Institutions,
and Markets

Led by IFPRI

IFPRI Discussion Paper 02065

December 2021

Enumerator Bias in Yield Measurement

A Comparison of Harvest versus Allometric Measurement of Coffee Yields

Vivian Hoffmann

Mike Murphy

Ezra Rwakazooba

Charles Angebault

Godfrey Kagezi

Giulia Zane

Markets, Trade, and Institutions Division

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

The International Food Policy Research Institute (IFPRI), a CGIAR Research Center established in 1975, provides research-based policy solutions to sustainably reduce poverty and end hunger and malnutrition. IFPRI's strategic research aims to foster a climate-resilient and sustainable food supply; promote healthy diets and nutrition for all; build inclusive and efficient markets, trade systems, and food industries; transform agricultural and rural economies; and strengthen institutions and governance. Gender is integrated in all the Institute's work. Partnerships, communications, capacity strengthening, and data and knowledge management are essential components to translate IFPRI's research from action to impact. The Institute's regional and country programs play a critical role in responding to demand for food policy research and in delivering holistic support for country-led development. IFPRI collaborates with partners around the world.

AUTHORS

Vivian Hoffmann (v.hoffmann@cgiar.org) is as Senior Research Fellow in the Markets, Trade, and Institutions Division of the International Food Policy Research Institute (IFPRI), based in Ottawa, Canada.

Mike Murphy* (m.murphy@cgiar.org) is a Senior Research Analyst in IFPRI's Markets, Trade, and Institutions Division, Washington, DC.

Ezra Rwakazooba (info@ignosi.org) is a Managing Partner and Data Scientist at Ignosi Research, based in Kampala, Uganda.

Charles Angebault (charles.a@metajua.com) is Managing Director at Metajua (U) Ltd, based in Kampala, Uganda.

Godfrey Kagezi (gkagezi@gmail.com) is Programme Leader, Coffee Cocoa Plant Health Management Programme (CPhM-p) at the National Coffee Research Institute (NaCORI), based in Mukono, Uganda

Giulia Zane (giulia.zane@gmail.com) is an Impact Evaluation Consultant at the Inter-American Development Bank.

**Corresponding author*

Notices

¹ IFPRI Discussion Papers contain preliminary material and research results and are circulated in order to stimulate discussion and critical comment. They have not been subject to a formal external review via IFPRI's Publications Review Committee. Any opinions stated herein are those of the author(s) and are not necessarily representative of or endorsed by IFPRI.

² The boundaries and names shown and the designations used on the map(s) herein do not imply official endorsement or acceptance by the International Food Policy Research Institute (IFPRI) or its partners and contributors.

³ Copyright remains with the authors. The authors are free to proceed, without further IFPRI permission, to publish this paper, or any revised version of it, in outlets such as journals, books, and other publications.

Abstract

Measuring yield accurately is critical for evaluating the impact of interventions that aim to increase agricultural productivity but presents challenges in the case of coffee due to the long harvest period. An allometric approach, in which the fruits on randomly selected branches and clusters are counted is widely used due to its non-destructive nature and acceptability to farmers. However, this approach requires careful attention to detail, which may be difficult to maintain in the context of large-scale data collection efforts. Using data from 199 small-scale Robusta coffee farms in Uganda, we compare yield estimates obtained through a standard allometric protocol against those from a one-time harvest of both ripe and unripe cherries prior to the start of the harvest season. The one-time harvest method was widely acceptable to farmers. Allometric yield estimates explain just under half of the variation in the harvest-based yield measure. While estimated yield is similar across methods for the first tree harvested per farm, we observe a larger difference in allometric versus harvest-based estimates, and systematically lower counts of stems and branches for trees assessed later during the farm visit. We interpret these findings as evidence of deteriorating enumerator performance on the allometric method over time, implying a risk of downward-biased yield estimates.

Keywords: agricultural training, impact evaluation, methods

Acknowledgements

This research was supported by Stiftung Coffee Agronomy Training, and the CGIAR Research Program on Policies, Institutions, and Markets (PIM), led by IFPRI. We are grateful to the data collection team for their hard work on this project.

Introduction

Measuring yield accurately is critical for evaluating the impact of interventions that aim to increase agricultural productivity. Farmer reports of yield may be highly variable due to both recall error and inaccurate plot size estimates (Gollin and Udry, 2021). They may also be affected by social desirability bias, especially in the context of interventions that aim to increase yield. For these reasons, weighing the harvest obtained from a randomly selected sub-plot is generally considered the most accurate way of measuring yield.

In the case of coffee, the harvest season spans several months, with coffee fruits (cherries) ripening at different points in time, even within the same tree (Da Matta et al., 2007). The most accurate way to measure coffee yield is thus through multiple harvests of whole selected plants over the course of a fruiting season. However, given the high cost of multiple farm visits, a more common approach is to multiply one-time pre-harvest counts or estimates of cherries per plant by the average weight of the ripe fruits or beans (Upreti et al., 1991; Soto-Pinto et al., 2000; Cilas and Descroix, 2009; Idol and Youkhana, 2019). This method, referred to as allometric measurement, has the additional advantage of being non-destructive and is thus acceptable to most farmers.

One potential shortcoming of allometric measurement is that accurately counting cherries requires careful attention to detail. Enumerators' attention may wane over time, reducing count accuracy. Especially if performance cannot be directly monitored, enumerators may guess at the number of branches, clusters, or cherries on a tree, rather than count them. Both fatigue and shirking are expected to increase measurement error and may, in addition, introduce bias of unknown direction.

A third yield measurement option that combines these two approaches is to do a one-time harvest of ripe and unripe fruit, sort these into categories by size, and report per-category weights (Campanha et al., 2004). Like the multiple yield-based method, this approach has the advantage of being less affected by potential enumerator attention constraints. However, this method has two potential drawbacks. First, stripping coffee trees of all leaves and cherries is known to increase susceptibility to disease and reduce harvest in subsequent years (Willson, 1999 as cited by Daniels, 2009). While a one-time harvest can be done more carefully and selectively than full stripping, the long-term yield impact of this practice is not known and farmers must thus be compensated based on potential future yield losses. Second, the acceptability to farmers of picking cherries ahead of the harvest is not known. In certain contexts, for example the region of western Uganda where data for this study were collected, some farmers sell the right to harvest coffee from their plots, rather than harvesting the crop themselves. Farmers may be

reluctant to allow coffee governed by such an arrangement to be harvested, as this could lead to conflict with the buyer.

In this paper, we use data from a sample of 199 coffee farms in Uganda to test the acceptability of one-time harvest yield estimation methodology and compare yield estimates obtained through this method against a standard allometric coffee yield measurement protocol. We generate multiple plant-level estimates at each site, in order to evaluate the suitability of the allometric procedure for generating farm-level yields.

1. Materials and methods

1.1 Sample selection

The study sample consists of coffee farmers in 22 coffee growing villages of Mateete subcounty within Ssembabule district, in Western Uganda. Villages were randomly selected from a list of villages identified by a non-governmental organization (NGO) as appropriate for inclusion in a coffee agronomy training program. The NGO had not conducted any intervention in the study villages at the time of data collection.

In each selected village, a listing of coffee farmers was conducted with the help of community leaders. Ten farmers were then randomly selected per village for inclusion in the study, and informed consent was obtained to inspect the coffee farm. In some villages, the number of coffee farmers available were fewer than 10, resulting in a sample size of 213 farmers. Of these, we exclude 14 cases for which one or both of the trees sampled for harvest-based measurement had a recorded yield of zero, resulting in an analysis sample of 199 farms.

1.2 Data collection

Data was collected during March and April of 2018, prior to the beginning of the main coffee harvest in the region, by enumerators recruited through a Ugandan research firm. Enumerators were paid a daily wage regardless of the number of households surveyed per day. The expectation was that each team of two enumerators would complete data collection on two to three farms per day, but the completion rate was closer to two per team on average, exceeding two farms on 10 days and falling short of this target on five days. Team leaders were told to prioritize quality over speed, and there were no ramifications for failing to miss the two-farm target. The following protocol was used to select coffee trees.

First, an enumerator measured the area of the coffee field using a hand-held GPS unit. If the area of the field was above half an acre, coffee trees on four separate 10 x 10 meter sampling squares were inspected.

For smaller fields, only three squares were used.

Next, the enumerator walked the longest diagonal connecting two opposite corners of the field, counting his or her steps. Three or four random stopping points were drawn by the data collection program, depending on the size of the field. Each stopping point constituted the northwest corner of a sampling square. Allometric yield measurement was conducted for the trees closest to the northwest and southeast corners of the square.

Consent to harvest the ripe and unripe cherries from two coffee trees, in exchange for monetary compensation, was obtained separately from consent to participate in the study.¹ All farmers except one agreed to allow the researchers to harvest their trees. For those who agreed, all cherries of at least 10 mm length on the trees closest to the northwest corner in the first two sampling squares were harvested. This threshold was used as smaller cherries were not expected to reach maturity within the coming harvest season (and may not reach maturity at all). As the ultimate goal of the yield estimation method tested was to measure the yield impact of a coffee agronomy training program, the conservative approach of omitting cherries that may or may not eventually be harvested was deemed preferable. Including cherries smaller than 10 mm in length could upwardly bias the estimated impact on yield if practices encouraged by the training led trees to produce more cherries that did not make it to harvest.

1.3 Allometric yield measurement

For the selected trees, allometric yield measurement was conducted to estimate the number of cherries (≥ 10 mm) by counting the number of main stems, and randomly selected subsets of branches, glomerules (cherry clusters) and individual cherries. We follow the procedure of Cilas and Descroix (2009) in sampling one main stem, four branches on that stem, and two clusters per branch. The number of cherries in each selected cluster were then counted.

Sampling was restricted to productive stems, defined as those with at least one cherry meeting our minimum length standard of 10 mm. Non-productive stems were not counted. Stems were labeled with papers indicating numbers from one up to the total number of stems, and a time-stamped photograph was taken. This was done to avoid the risk of enumerators purposively selecting stems with fewer cherries. Next, the number of branches on the stem (whether productive or not) were counted. This number was entered into a tablet computer, and four branches were randomly selected by the data collection

¹ Compensation was set 50 percent higher than the current price of coffee cherry, to compensate for the fact that the coffee had not yet reached maturity, as well as the potential negative impact on future yield.

program. Clusters of cherries on each of the selected branches were then counted, this number entered, and two clusters per branch were selected. Finally, the number of cherries (≥ 10 mm in length) on each of the randomly selected clusters were counted.

1.4 'One-time' harvest-based yield measurement

Two of the trees that had been measured allometrically were fully harvested by enumerators, immediately following allometric measurement. Cherries under 10 mm were discarded, while the remainder were grouped into three categories based on length: 10 to 12 mm, 12 to 14 mm, and those longer than 14 mm. The share of all cherries in each size group was calculated. Ripe (red) cherries were then separated from the unripe (green) cherries and sorted into the same size groups.

2. Analysis

2.1 Estimation of yield per tree

For both methods, the average weight of a sample of (both unripe and ripe) cherries for each size category was used in order to estimate the overall yield. For the harvest method, the total harvested weight of cherries in each category was divided by the mean measured weight of cherries in that category to obtain an estimate of the number of cherries on the tree in each size category. These values were then summed across categories to give an estimate of the total number of cherries (≥ 10 mm) on the tree.

For the allometric estimate, the number of cherries counted per cluster was multiplied by the total number of clusters counted on sampled branches, the total number of branches counted on sampled stems, and the total number of stems counted on the tree.

Finally, for both the harvest-based and allometric methods, the estimated yield was calculated according to the following formula:

$$\text{total cherries} \cdot \sum_{i=1}^3 \text{share of ripe cherries in size category } i \\ \cdot \text{weight of a ripe category } i \text{ cherry}$$

Differences between the yield estimates given by the two approaches are therefore determined by differences in the estimated number of ripe cherries (within each category).

2.2 Comparison of harvest-based versus allometric yield estimates

Table 1 presents summary statistics for coffee yields using the two methods, excluding trees with a harvest-based yield of zero.

Table 1. Summary statistics for yields estimated using harvest and allometric methods

<i>Estimation method</i>	Mean (Arithmetic)	SD	Min.	25th percentile	Median	75th percentile	Max.	N
Harvest	1.87	1.94	0.04	0.59	1.23	2.58	16.07	199
Allometric	2.24	3.67	0.02	0.46	1.12	2.58	35.62	199

The allometric method has a higher variance than the weight-based method (which is to be expected given the use of within-tree sampling of stems, branches, and clusters in the allometric method to determine cherry count). While the allometric method gives similar (if slightly lower) estimates near the center of the distribution, there is substantial divergence in the tails, resulting in a much higher estimated maximum yield (Figure 1). As the distributions obtained using both methods are highly skewed, we use the geometric mean and lognormal distributions when comparing the two measures.

Figure 1. Distribution of yield estimates based on harvest and allometric methods (logarithmic scale)

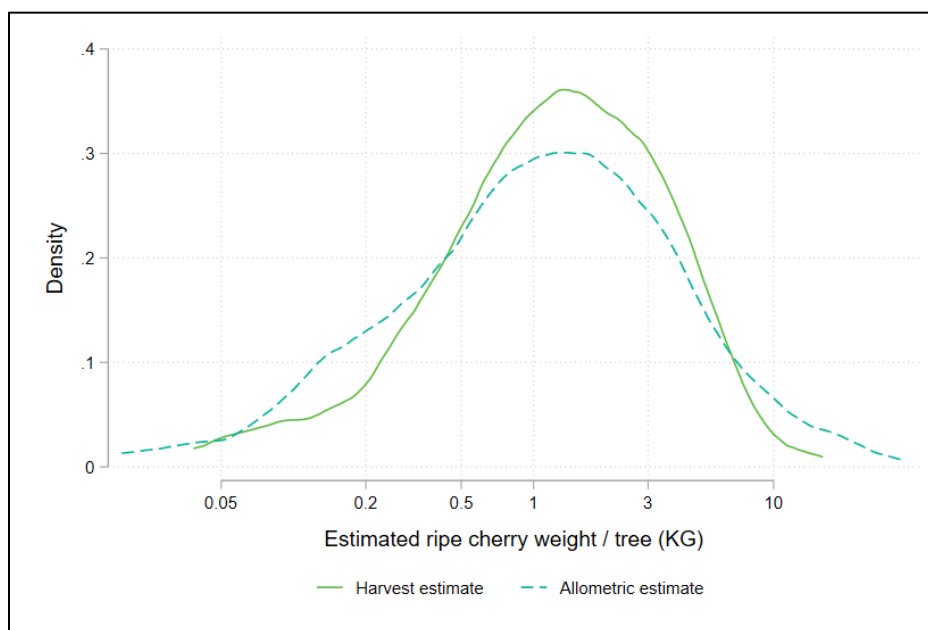


Table 2 presents the geometric mean of yield per tree using each estimation method, differences in these means, and p-values testing the equality of yield distributions across methods. In addition to comparing the mean yield for the two trees assessed on each farm, we also present comparisons for each of the trees

assessed using both harvest and allometric methods (first and third visually inspected).

Table 2. Comparison of geometric means using harvest and allometric methods (KG)

	Harvest	Allometric	Difference in means	Equality of distributions p-value ¹	Observations
Both trees (averaged)	1.14	1.02	0.13	0.325	199
First tree only	0.81	0.69	0.13	0.390	199
Second tree only	0.79	0.60	0.19	0.051	199

¹ Kolmogorov-Smirnov test of equality of distributions

Overall, the harvest-based estimate (GM = 1.14) is higher than the allometric measurement (GM = 1.02), though the difference in distributions is not statistically significant (p=0.325). The results are similar when comparing measurements of the first tree assessed, for which the geometric mean values for the harvest and allometric methods are 0.81 and 0.69 KG respectively (difference = 0.13, comparison of distributions p=0.390). There appears to be some deterioration in allometric performance for the second tree, for which the difference in means is greater (0.19 KG) and the difference in distributions is marginally significant (p=0.051). In Section 3.3 below, we explore the possibility of whether enumerator performance declines with time by examining differences in the number of tree components measured on the first tree measured per farm versus those assessed later.

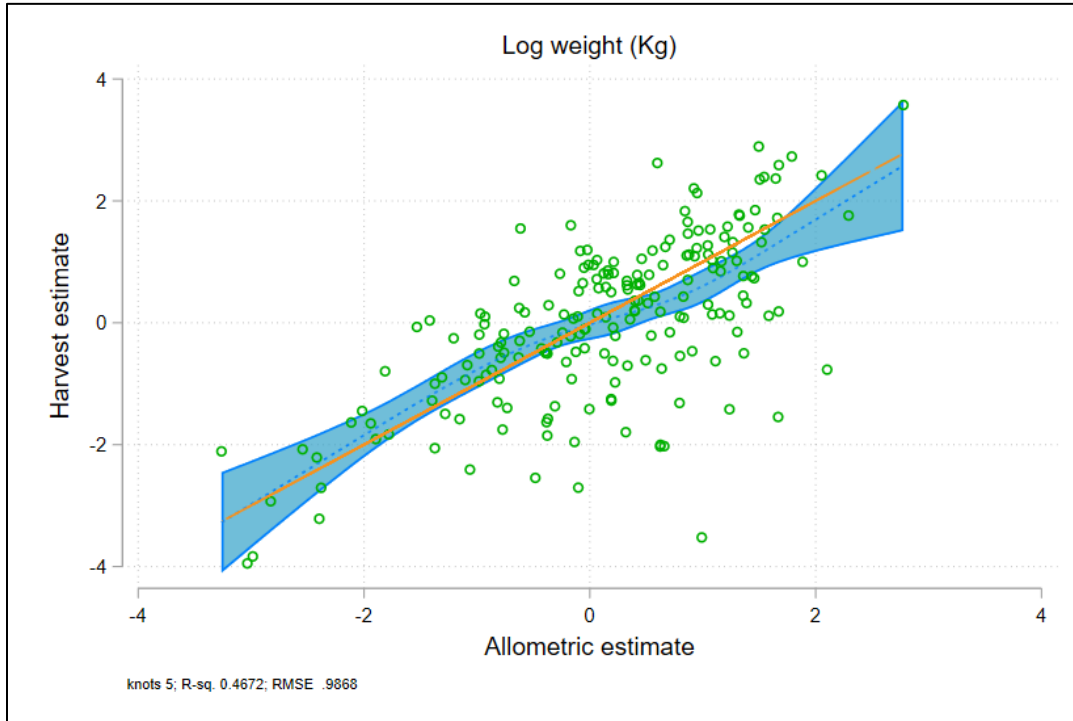
Table 3 uses a regression approach to assess the performance of allometric measurement in predicting the yield obtained from the harvest method, based on both aggregate (mean of both trees on each farm measured using the two approaches) and per-tree yield measurements. Across these, the allometric measure explains roughly half of the observed variation in harvested coffee weight (with an R² value in the range 0.45-0.48).

Table 3. Regression of harvest measurement on allometric measurement

Measurement	Dependent variable: Harvest measurement (Log)		
	Aggregate (1)	1st tree only (2)	2nd tree only (3)
Allometric measurement (Log)	0.557*** (0.043)	0.603*** (0.054)	0.578*** (0.037)
R-squared	0.458	0.478	0.453
Observations	199	199	199

Figure 2 plots the data from the two measures as a scatter plot (hollow circles), along with a line showing the ideal fit (45-degree solid line). We overlay a smoothed cubic spline (dashed line) and associated confidence interval.

Figure 2. Observed harvest values with restricted cubic spline (95% CI) for allometric estimate



The degree of dispersion in the allometric measure is much greater for the respective tails of the distribution, even when using a log scale. This is reflected in the much higher coefficients of variation for the top & bottom deciles (0.54 and 0.46 respectively, compared to a maximum value of 0.15 for other deciles). This greater degree of dispersion both results in relatively poor performance of the allometric measure for the tails of the distribution and decreases the overall predictive accuracy of the regression estimate generated using the allometric measure.

2.3 Performance of the allometric measure across tree measurements

One potential concern in employing the allometric approach is that enumerator performance may decline due to fatigue (lower count accuracy) or shirking (non-random selection of trees with fewer stems, branches, clusters or cherries to be counted; guessing rather than counting tree components) in order to finish faster. Purposive selection of smaller, easier to measure trees, is also a potential concern for both approaches. To explore this, in Table 4 we compare the mean value and standard deviation of each of the

components measured for the first versus sixth, and first versus eighth, tree counted.²

Table 4. Means and standard deviations of allometric measurement components by tree order

		Tree 1	Tree 6	p-value	Tree 1*	Tree 8	p-value
Allometric Estimate	Mean	1,787	1,754	0.898	1,920	1,499	0.191
	SD	2,842	2,534	0.115	2,808	2,675	0.573
Stems (per tree)	Mean	4.12	3.71	0.129	4.10	2.66	0.000
	SD	3.25	2.63	0.003	3.09	2.84	0.265
Branches (selected stem)	Mean	14.56	13.06	0.130	15.48	8.92	0.000
	SD	10.71	10.21	0.501	10.96	9.93	0.204
Clusters (mean per branch)	Mean	3.80	3.75	0.789	3.95	3.61	0.228
	SD	2.24	2.00	0.120	2.31	2.44	0.520
Cherries (mean per cluster)	Mean	6.02	5.71	0.300	6.24	5.56	0.057
	SD	3.22	3.43	0.379	3.17	2.74	0.095
Observations		199			136		

* Sample restricted to farms where observations are available for Tree 8. p-values for comparisons of means are calculated using Poisson regression.

The estimated mean for the eighth tree using the allometric method is 17% lower than that for the first tree, though given the high variability of allometric yield estimates, the difference is not statistically significant. Decomposing by the respective counts of each part, the number of components observed for the sixth tree are uniformly lower for all components than for the first tree (again the differences are not statistically significant). Among the smaller sample of larger farms on which eight trees were counted, we do observe statistically significant differences in the number of stems, cherries and branches counted between the first and eighth tree, with counts lower for the later tree. In general, we do not observe significant differences in the dispersion of the respective counts, with the exception of the comparison of stems between the first and sixth trees. Our findings suggest the existence of a trade-off to counting additional trees, with poorer enumerator performance potentially offsetting reductions in sampling variance.

² Since the number of trees counted varied by farm size, we compare both the first and sixth measures since these are available for all farms, as well as the first and eighth, which restricts the sample to large farms for which the latter is available.

3. Discussion

We tested a one-time harvest method for estimation of the Robusta coffee yield in smallholder farms in western Uganda. This method was widely acceptable to farmers; only one of 213 farmers declined to allow researchers to harvest all cherries measuring at least 10 mm from two of their coffee trees, in exchange for compensation.

Our method is similar to that of Campanha and colleagues, but we take the additional step of estimating the number of cherries harvested, to allow comparability with allometric results. We find that a standard allometric protocol produces slightly lower, but statistically indistinguishable estimates of yield compared to our harvest-based measure for the first and third tree measured. Allometric estimates are poorer for very high and low values of the harvest estimate, where there is greater relative dispersion in the observed data than for other portions of the data.

The gap in allometric versus harvest-based estimates increases with tree sequence, and the number of stems and branches reported for trees inspected later are significantly lower, suggesting that enumerator error increases with workload. Enumerators for this project were closely monitored, and it seems reasonable to conclude that our estimates of enumerator bias represent a lower bound of what might be expected in a larger-scale survey effort, where monitoring intensity is lower. Ensuring that enumerators are able to rest and address their physical needs during the workday, for example by requiring that firms make it clear that employees have a right to breaks, and provide them with a packed lunch, could be investigated as a way to curb fatigue.

To the extent that the lower number of stems and branches on later trees is driven enumerators' purposive selection of smaller trees to avoid workload, this source of bias may also be a problem for yield measurements based on harvest of a sample of trees. One approach to mitigating this source of bias would be to have a separate team select and mark trees for harvesting. Photographs of these trees taken by both teams could be used to monitor the harvest team's compliance to study protocols.

Noting the potential negative impact of harvesting unripe coffee cherries on future yields, the number of trees to which the one-time harvest-based method of yield measurement tested in this paper is applied should be kept to a minimum, and farmers should be compensated for the potential damage to their trees beyond the value of the coffee harvested. Future work could assess the impact on yields of this method over multiple harvests to determine the appropriate level of compensation. In the absence of such evidence, an upper bound of the potential yield impact should be used.

References

- Bote, A.D., Vos J. 2016. Branch growth dynamics, photosynthesis, yield and bean size distribution in response to fruit load manipulation in coffee trees. *Trees* 30(4):1275-1285.
- Campanha, M.M., Santos, R.H.S., De Freitas, G.B., Martinez, H.E.P., Garcia, S.L.R. and Finger, F.L., 2004. Growth and yield of coffee plants in agroforestry and monoculture systems in Minas Gerais, Brazil. *Agroforestry systems*, 63(1), pp.75-82.
- Cilas, C. and Descroix, F. 2009. Yield estimation and harvest period. In: Wintgens JN (ed) *Coffee: growing, processing, sustainable production. A guidebook for growers, traders, and researchers*, 2nd edition. Wiley, Weinheim, pp 601–609.
- Da Matta, F.M., Ronchi, C.P., Maestri, M., Barros, R.S. 2007. Ecophysiology of coffee growth and production. *Brazilian Journal of Plant Physiology* 19(4):485-510.
- Daniels, N., 2009. *Variations in coffee processing and their impact on quality and consistency* (M.Sc. thesis, Michigan Technological University). 80 pp.
- Gollin, Douglas, and Christopher Udry. 2021. Heterogeneity, measurement error, and misallocation: Evidence from African agriculture. *Journal of Political Economy* 129(1): 1-80.
- Idol T.W. and Youkhana AH. 2019. A rapid visual estimation of fruits per lateral to predict coffee yield in Hawaii. *Agroforest Systems* pp. 1-13. <https://doi.org/10.1007/s10457-019-00370-y>.
- Soto-Pinto, L., Perfecto, I., Castillo-Hernandez, J. and Caballero-Nieto, J., 2000. Shade effect on coffee production at the northern Tzeltal zone of the state of Chiapas, Mexico. *Agriculture, ecosystems & environment*, 80(1-2), pp.61-69.
- Upreti, G., Bittenbender, H., Ingamells, J.L. Rapid estimation of coffee yield. *Quatorzieme colloque scientifique international sur le café*. pp. 585-593.
- Willson, K. C., 1999. *Coffee, Cocoa, and Tea*. CABI Publishing, New York, 300p.

ALL IFPRI DISCUSSION PAPERS

All discussion papers are available [here](#)

They can be downloaded free of charge

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

www.ifpri.org

IFPRI HEADQUARTERS

1201 Eye Street, NW
Washington, DC 20005 USA
Tel.: +1-202-862-5600
Fax: +1-202-862-5606
Email: ifpri@cgiar.org