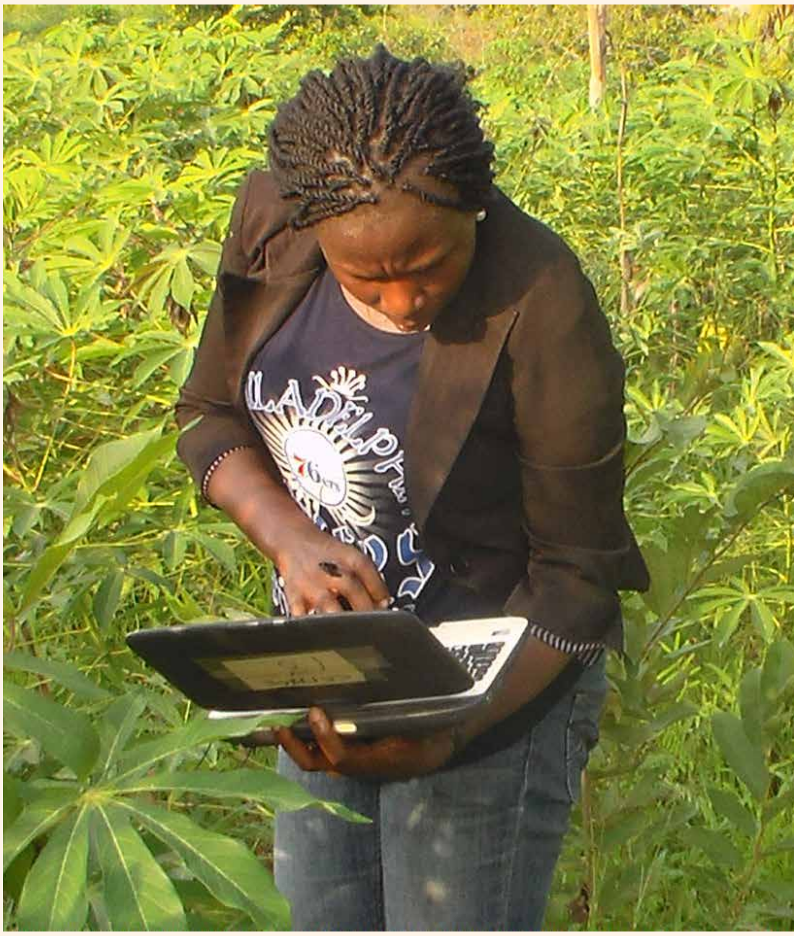
IITA
Transforming African Agriculture

CGIAR
RESEARCH PROGRAM ON
Roots, Tubers and Bananas

The Cassava Monitoring Survey in Nigeria

Tesfamicheal Wossen, Gezahegn Girma, Tahirou Abdoulaye, Ismail Rabbi, Adetunji Olanrewaju, Arega Alene, Shiferaw Feleke, Peter Kulakow, Godwin Asumugha, Adebayo Abass, and Victor Manyong



www.iita.org

The Cassava Monitoring Survey in Nigeria

Tesfamicheal Wossen, Gezahegn Girma,
Tahirou Abdoulaye, Ismail Rabbi, Adetunji Olanrewaju,
Arega Alene, Shiferaw Feleke, Peter Kulakow,
Godwin Asumugha, Adebayo Abass, and Victor Manyong

International Institute of Tropical Agriculture, Ibadan.

February 2017

Published by the International Institute of Tropical Agriculture (IITA)
Ibadan, Nigeria. 2017

IITA is a non-profit institution that generates agricultural innovations to meet Africa's most pressing challenges of hunger, malnutrition, poverty, and natural resource degradation. Working with various partners across sub-Saharan Africa, we improve livelihoods, enhance food and nutrition security, increase employment, and preserve natural resource integrity. It is a member of the CGIAR System Organization, a global research partnership for a food secure future.

International address:
IITA, Grosvenor House,
125 High Street
Croydon CR0 9XP, UK

Headquarters:
PMB 5320, Oyo Road
Ibadan, Oyo State

ISBN 978-978-8444-81-7

Correct citation: Wossen, T., G. Tessema., T. Abdoulaye, I. Rabbi, A. Olanrewaju, A. Alene, S. Feleke, P. Kulakow, G. Asumugha, A. Adebayo, and V. Manyong. 2017. The cassava monitoring survey in Nigeria final report. IITA, Ibadan, Nigeria. ISBN 978-978-8444-81-7. 66 pp.

Printed in Nigeria by IITA

Cover photo: A survey enumerator measuring a cassava field with handheld Global Positioning System and recording data into a computer tablet.

Contents

Tables

Figures

Acronyms

ADP	Agricultural Development Project/ Agricultural Development Program
BNARDA	Benue State Agricultural and Rural Development Authority
CMS	Cassava Monitoring Survey
DNA	Deoxyribonucleic acid
FGD	Focus group discussion
GBS	Genotyping by sequencing
GDF	Genomic diversity facility
GPS	Global positioning system
IBS	Identity by state
IITA	International Institute of Tropical Agriculture
LGA	Local government area
MAF	Minor allele frequency
M&E	Monitoring and evaluation
NARS	National agricultural research systems
ng/µl	Nano gram per micro liter
NRCRI	National Root Crops Research Institute
OAU	Obafemi Awolowo University
QC	Quality control
RTB	CGIAR Research Program on Roots, Tubers and Bananas
SNP	Single nucleotide polymorphism
TME	Tropical *Manihot esculenta*. A group of early improved varieties
TMS	Tropical Manihot species

Acknowledgements

Funding for this study came from the CGIAR Research Program on Roots, Tubers and Bananas (RTB), and the Bill & Melinda Gates Foundation (BMGF) through the Cassava Monitoring Survey CMS project at IITA. The authors would like to thank Dr Graham Thiele, Director of RTB, Dr Laurence Kent, and other Gates Foundation staff (past and present) for their guidance and encouragement during the design and implementation of CMS. They especially thank Jared Jones (former Gates Foundation) and Elvis Fraser of Gates Foundation for their support and guidance throughout the implementation of the CMS project. The contributions of all the participants of the Nigeria Cassava Monitoring Convening Workshop organized, 15–21 March, 2015 by BMGF in Dar es Salam, Tanzania are gratefully acknowledged. The contributions of Dr Alfred Dixon, Dr Holger Kirscht (former IITA staff), Mr Henry Musa Kpaka to the conceptualization and design of the CMS project are also gratefully acknowledged. The contribution of the National Roots Crops Research Institute, Umudike, Nigeria and other local partners to the success of the CMS Project is also acknowledged. The enumerators, extension agents, and IITA staff from socioeconomics, bioscience and cassava breeding units are recognized and thanked for their contributions. A special thanks to all the farmers and other people who generously gave their time for the interviews.

Executive summary

The Bill & Melinda Gates Foundation (BMGF) and the CGIAR Research Program on Roots, Tubers and Bananas (RTB) cofunded the Cassava Monitoring Survey (CMS) project with the objective of documenting the extent of adoption of improved cassava varieties with a rigorous sampling and DNA based identification of improved varieties in Nigeria. For examining the determinants of adoption, the project undertook the following four major tasks: (i) Large-scale household (HH) survey (2500 households from the states responsible for 80% of cassava production in Nigeria), (ii) DNA-based varietal identification using single nucleotide polymorphism (SNP), (iii) GPS-assisted area measurement and (iv) gender-differentiated end-user surveys on varietal and trait preferences. This Executive summary presents the main results from each component.

i) Results from large-scale household survey

Results from CMS showed that 60% of farm households have adopted improved cassava varieties in Nigeria. The distribution of adoption by gender also revealed that the adoption rate among male-headed households (MHHs) was about 61.5% while among female-headed households (FHHs) it was relatively low, about 48.6%. The adoption rate across the different regions of the country showed a large spatial heterogeneity. In particular, adoption rates were as high as 79% in the Southwest region but only about 31% in the Southeast. In terms of preference heterogeneity, more than 70% of the farmers considered varietal traits such as quality of garri or fufu*/akpu*, high root yield, big root size, high market demand, and early maturity as very important. In addition, a substantial share of respondents considered traits such as storage, white root color, and pest and disease resistance as very important. However, more than 40% of the farmers reported varietal traits such as yellow root color, root shape, starch content, and good tapioca taste as not important. Based on farmers' self-reported data, the main determinants of adoption can be categorized into the following groups:

(i) **Information market imperfections:** Both extension access and mobile ownership positively affected the probability of adopting improved varieties

(ii) **Credit market imperfections:** Both access to formal credit sources and membership of social networks that relax credit market constraints such as membership in informal Savings and Credit associations as well as membership in cooperatives affected the probability of adopting improved varieties positively.

(iii) **Labor market imperfections:** Households with a large size are more likely to adopt improved varieties, suggesting that availability of labor affects the probability of adoption positively.

(iv) **Heterogeneity in plot fertility and management:** We found that soil fertility and plot management are important determinants of adoption. In particular, farmers are less likely to grow cassava on plots of medium fertility compared with highly fertile plots. Moreover, farmers are more likely to adopt improved varieties on plots managed by men or jointly by men and women compared with the plots managed by women.

(v) **Trait preference heterogeneity:** Individual preference heterogeneity in improved varieties affected adoption decisions. For instance, farmers who perceived traits such as quality of garri, root yield, and early maturity as highly important are more likely to adopt improved varieties. Surprisingly, we found a significant and negative effect for the trait, starch content. This might be due to the lack of local processors for starch. The major determinants of disadoption included the following: distance to market, lack of planting material, availability of better improved varieties, and pest and disease problems.

ii) Results from DNA-based varietal identification using SNPs

While combining DNA fingerprinted data with the CMS socioeconomic data, average adoption rates tended to be similar. Even though the average adoption rates between farmers' self-reported and DNA fingerprinted data seem similar, the misclassification rate is large. Misclassification happens when farmers who think they are growing improved varieties actually grow local varieties or when farmers who think they are growing local varieties actually grow improved varieties. Therefore, for examining the determinants of adoption, the measurement of "improved varieties" matters as the determinants of adoption are different while using farmers' self-reported and DNA fingerprinted data. Three sources of heterogeneity largely explain the probability of correctly classifying cassava into improved and local varieties. These include level of education, access to information (such as mobile phone ownership and access to extension access), membership of formal and informal organizations, and location.

iii) Results from Global Positioning System (GPS)-assisted area measurement

In terms of GPS-based cassava plot measurement, we found that GPS-based area measurements and those based on self-report from farmers produce different results. In particular, farmers overestimated the size of small farms and underestimated the size of large farms. The measurement error (the discrepancy between GPS-assisted and self-reported areas) is normally distributed.

iv) Gender-differentiated end-user surveys on varietal and trait preferences

Our findings from focal group discussions confirm that farmers like varieties that are high-yielding (with many big roots), especially if the cassava tolerates poor soil, early maturing (but they may also need some that are late maturing for food security (especially if late maturing cassava stores better underground) and varieties that store underground for a long time. Even though regional heterogeneities are large, we found traits that are common in all regions. Thus, farmers want high yielding, early maturing cassava that stores well underground. The Southwest expressed a need for cassava that controls weeds, and the North wanted cassava with resistance to mealybugs. The South-South wants varieties that tolerate poor soil. In addition, we also found important g***ender differences*** in trait preferences***.*** In particular, women want cassava that is easy to peel. This is by far the most important gender difference. In the Southwest and the North, men asked for markets for cassava roots but women had little or no trouble in selling their finished products, such as garri. Both women and men in all regions were eager to try improved varieties. However, farmers were not using more improved varieties because of a lack of access to planting material. If improved planting material were more widely available, farmers would experiment with it and probably adopt many varieties. Disadoption occurs for local varieties***.*** Occasionally an early improved variety has been disadopted, after 20 or 30 years.

Furthermore, our gender-disaggregated econometric analysis on data from spouse and household head revealed that preference trait heterogeneity between the spouse and household head affects adoption decisions differently. We found that important traits such as early maturity, quality of fufu and garri, root yield and storability have statistically significant effects on the adoption decision of both the spouses and household heads. However, traits such as stem yield, poundability, and pest and disease resistance affected the adoption decision only of spouses whereas the trait, ease of loading, affected the adoption decision only of household heads. These results underscore the fact that asking the spouse and household head the same question could result in different determining factors of adoption. For some traits such as tapioca taste, effects were positive and significant when using responses from spouses but negative and significant when using responses from the household heads. Consequently studies that assume a unitary (male) decision-maker as the basis of their analysis may miss some underlying factors. The importance of each factor on the adoption decision of the household head/spouse depends on their individual relative importance in agriculture related activities.

Project milestones

General introduction

The Cassava Monitoring Survey in Nigeria (CMS) was designed to assess the adoption of improved cassava cultivars and reveal the drivers of adoption and disadoption in Nigeria, the largest producer of cassava in the world. Focusing on cassava, this report examines the following empirical questions.

i) What are the levels of adoption of improved cassava varieties?
ii) What are the factors driving adoption and disadoption of improved varieties of cassava?
iii) What are the preferences of different end users for varietal attributes in terms of production, processing, and consumption traits?
iv) Are there gender differences associated with varietal adoption, preferences and adoption pathways?

Understanding how and why households decide what varieties to adopt is crucial for ex-post impact assessment of adoption. In this project, we combined a large household (HH) survey, DNA-based varietal identification, GPS-based area measurement, and gender-disaggregated trait preferences to answer these policy relevant questions. The report gave special attention to individual preference heterogeneity for cassava traits as it affects adoption decisions through various channels (Useche et al. 2009; Useche et al. 2013; Bajari and Benkard 2005; Hensher et al. 1999; Ladd and Suvannunt 1976). New and improved cassava varieties could, for example, reduce cooking time or have a long/short growing period, and these considerations factor into the adoption decisions of farmers (Useche et al. 2009; Useche et al. 2013). Especially in the context of Nigeria, genetic improvements on cassava can affect the taste of edible cassava varieties, which can then affect farmer utility and thus adoption. As such, ceteris paribus, adoption decision of a new cassava variety may be different for two similar individual farmers as traits may differentially affect their food consumption habits. In this project, by explicitly accounting for trait preferences, we draw conclusions that help policymakers to develop targeting interventions for addressing supply-side constraints (for example, credit, infrastructure, or market failure, as well to identify the specific cassava varieties that fit the circumstances of farmers (demand-driven breeding) instead of advocating a “one size fits all” intervention (supply-driven breeding).

In an attempt to circumvent traditional survey-based measurement errors in varietal identification and area measurement, this project further implemented two novel approaches: DNA-based varietal identification and GPS-based cassava plot measurements. Ex-post assessment of adoption rates and the benefits of adoption heavily rely on farm household surveys. However, the magnitude of measurement errors and self-reported bias in adoption surveys has never been taken into account, leaving the bias (direction) and standard errors (magnitude) of existing adoption studies unknown. In addition, the bias in the magnitude and direction of reported regression coefficients for supply-side constraints emanating from measurement errors is also widely unknown (Maredia & Reyes 2015). The main sources of measurement errors in self-reported adoption rates include seed contamination by dealers or farmers and inability of farmers (extension agents) to correctly identify what is an improved vs. a traditional variety. Irrespective of the source of the bias, the wrong identification of varieties creates a problem in identifying the appropriate treatment and counterfactual group for possible impact evaluation as well for correctly identifying the main determinants of adoption and disadoption.

In the presence of measurement errors (even after controlling for endogeneity), reported estimated effects capture the pure technological effect plus random/systematic measurement error effects (Manski 2004; Bulte et al. 2014; Chassang et al. 2015; Aldashev et al. 2016; Lewbel 2007; Hasselt and Bollinger 2012). The use of data from DNA-based varietal identification can produce credible point estimates for the ***pure technological effect*** by eliminating possible measurement and self-reported biases. Failure to address the problem of bias (when bias is systematic) may lead to wrong conclusions and policy interventions, especially when targeting is needed. Even worse is when heterogeneity effects are considered in terms of gender or assets along the income distribution. For instance, the effect of credit or access to extension on technology adoption could have a wrong sign and magnitude if the bias from self-reported adoption surveys is large enough. In addition, reported impacts on yield/income could well be wrong if improved varieties are wrongly recorded as traditional and vice versa.

However, there is a caveat in using DNA-based varietal identification as a benchmark due to the inherent behavioral adjustment of farmers based on their own subjective self-assessment of a variety (Bulte et al. 2014). This is the case as productivity of a certain variety depends not only on the technology itself but also on the use of complementary inputs such as fertilizer, labor, or irrigation (Bulte et al. 2014; Chassang et al. 2015). While using DNA-based varietal identification data, certain dimensions of the farmer's efforts are unobservable which leads to biased causal inference (De Janvry et al. 2011; Bulte et al. 2014). For instance, conditioning for all other confounding factors, a given farmer will certainly allocate more labor and fertilizer and apply better management techniques for an improved variety than a traditional variety. Since the level of efforts and field management of plots depends on a farmer's own perceptions and beliefs, caution is needed in using DNA fingerprinted results as a benchmark for impact assessment. In an effort to minimize the two sources of bias (self-reported from survey data and behavioural adjustment from DNA fingerprinted data), we used a unique data set that combines self-reported and DNA-based adoption rates to provide robust results on the determinants of adoption.

Another related measurement error is related to plot measurement. In the mainstream development economics literature the inverse relationship between farm size and productivity is taken as a stylized fact (Collier 1983; Van Zyl et al., 1995; Barrett 1996; Kimhi, 2006; Barrett et al. 2010; Carletto et al. 2013; Dillon et al. 2016; Holden &Fisher 2013). However, such relationships are mostly based on self-reported area measurement, which is prone to measurement errors. In many household surveys, farm sizes are measured by simply asking farmers to estimate the area of their farm plots. While this approach is simple and inexpensive (in terms of money and survey time), it may produce imprecise values for many reasons: Firstly, most farmers have limited understanding of the purposes of surveys. As such they may report wrong values if they perceive that the information collected by enumerators is used as a basis to take away their land as part of a large land re-distribution policy (Dillon et al. 2016; Holden and Fisher 2013). This is the case in the context of many developing countries as land tenure systems are quite fragile or non-existent. Secondly, due to a low level of formal education and underdeveloped land markets, farmers are often unfamiliar with standard units of area measurement (Dillon et al. 2016). Even when local units of measurement are used, conversion to standard area measurement units is also prone to errors. Therefore, estimates of farm size based on self-reporting could be considered inaccurate. However, these errors can be consequential as ex-post assessment of the benefits of adoption relies heavily on the accurate measurement of land size and production values.

In an effort to measure the level of measurement errors, the CMS project introduced Global Positioning System (GPS)- based area measurement in Nigeria. This serves to measure the magnitude of self-reported bias as well as to undertake a credible causal analysis of the determinants of adoption. This is particularly important as the discrepancy between self-reported and GPS-based area values may become substantial and in some cases may vary by farm size or level of formal education.

Data collection instruments and sampling strategy

Working with all the partners on the project, the CMS questionnaire and the Surveybe[1] programming were finalized in May 2015. Key project staff and survey enumerators had been recruited in April 2015 and were trained on the use of Surveybe software. The data collection instruments were pre-tested during the training of the enumerators 20–22 May 2015. Given the more technical requirements of the type of data to be collected for the CMS project, a strong emphasis was placed on the recruitment of enumerators with at least a first university degree (BSc/BA) to participate in data collection. Furthermore, the project's field materials were procured including a van for fieldwork, computer tablets, and others materials.

Developing the sampling frame

The list of enumeration areas (EAs) for conducting the national census in Nigeria was obtained from the National Population Commission of Nigeria (NPCN). The EA list was obtained for the 16 states that contribute at least 80% of the total production of cassava in Nigeria. The states cut across four geopolitical regions (Table 1). Relying on agricultural development programs (ADPs) in the targeted states, a prior visit was made to each of the selected EAs to develop the lists of all cassava growing households. This lists provided a sampling frame for the selection of at least 50 cassava growing households in each EA, of which five household heads and two spouses were interviewed. This exercise facilitated the unbiased selection of samples for the final interviews. Furthermore, this prior visit also provided an opportunity to brief the sampled EAs on the CMS project and to help the staff of the ADPs that would finally provide guidance to the CMS survey teams to locate the randomly selected EAs.

Table 1. CMS study regions and states in Nigeria.

s/n	Region	States
1	Southwest	Ogun
2	Southwest	Ondo
3	Southwest	Oyo
4	Southwest	Ekiti
5	Southwest	Osun
6	North	Kaduna
7	North	Nasarawa
8	North	Taraba
9	North	Benue
10	North	Kogi
11	Southeast	Enugu
12	Southeast	Imo
13	Southeast	Anambra
14	South-South	Cross River
15	South-South	Akwa Ibom
16	South-South	Delta

1 It is a type of computer-assisted personal interview software.

Sample selection

As mentioned earlier, this survey used a multistage clustering sampling design, stratified by four well-established geopolitical regions of Nigeria: Southeast, South-South, Southwest and North. Each region had a sample size of 125 EAs. Furthermore, from each EA, random samples of 50 cassava growing households were selected for interview, and five out of the 50 were finally interviewed. This gave a total of 625 households that were interviewed in each region. In line with the initial proposal, interviews were conducted with 2500 randomly selected cassava growing households in the major cassava growing states in the four geopolitical regions of Nigeria that together account for over 80% of the annual total cassava production (See Fig. 1). The initial plan was to restrict data collection to the states in the North Central only; however, because Kaduna State (in the Northwest) and Taraba State (in the Northeast) were evidently among the states that account for 80% of the total cassava production in Nigeria, they were included in the study. Their inclusion gave an opportunity to collect data whose analysis and results can be extrapolated to cover the entire northern region.

Training of survey enumerators and pretesting of data collection instruments

The CMS project organized a 3-day training workshop from 20 to 22 May 2015 for the survey enumerators that would carry out data collection. A total of 60 survey enumerators were trained on the campus of IITA in Ibadan and provided a pool of enumerators that were used for data collection. The training was one of the most important components of the prerequisites for the successful use of the data collection tool (i.e., Surveybe program) and the collection of high quality data. The training workshop

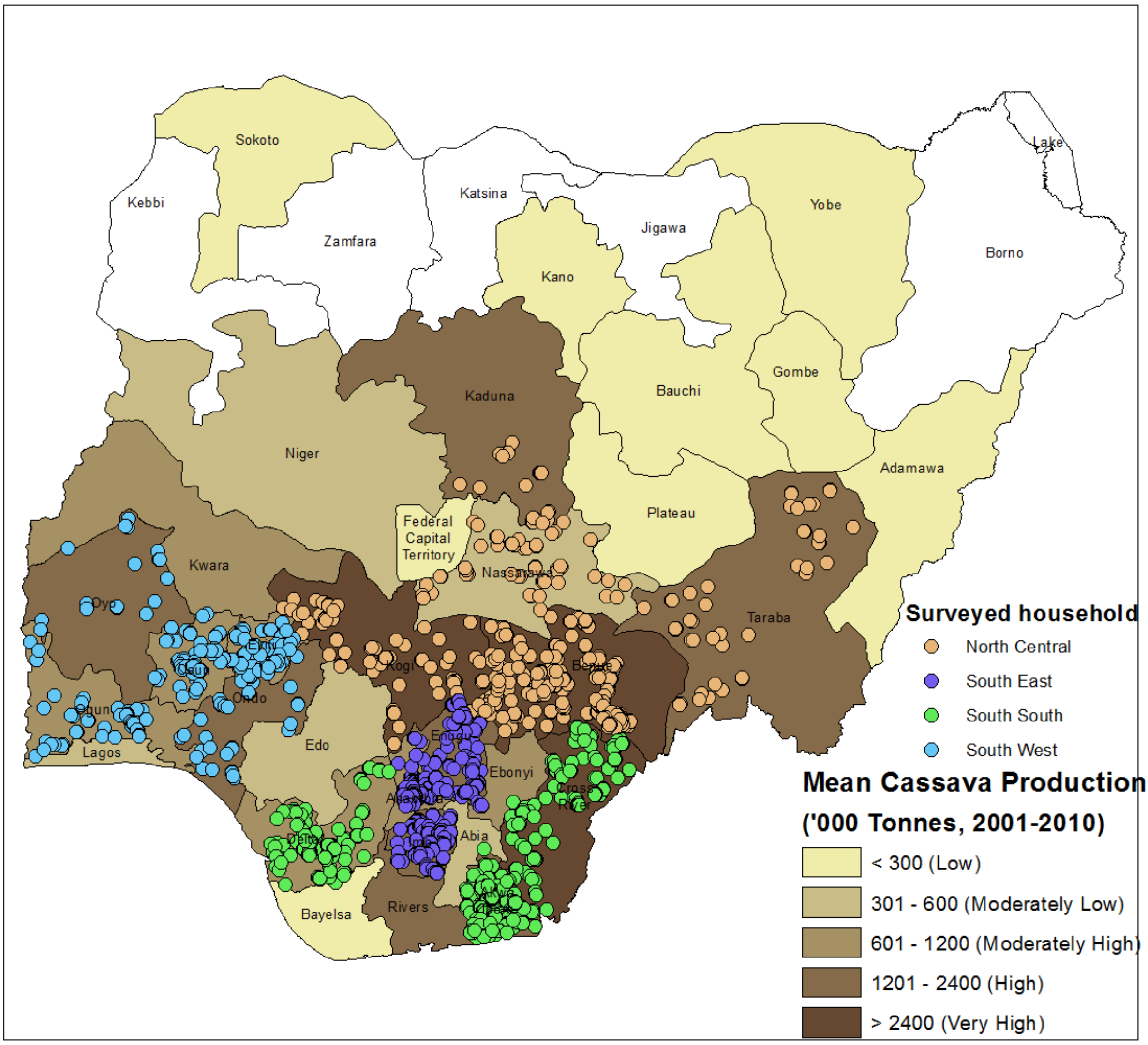


Figure 1. Study areas of CMS.

The points on the map represent the distribution of the HHs across the four study regions.

Picture 1. A cross-section of survey enumerators during the training on CMS research approach and the use of Computer Assisted Personal Interview tools.

equipped participants with practical skills on the use of Surveybe software for socioeconomic data collection, the methods of collection and preservation of cassava samples for DNA analysis, the use of GPS for area measurement, and training on the use of agronomic traits to differentiate the various cassava cultivars on-farm. The training participants undertook a one-day field visit to three cassava-producing communities around Ijaye town in Oyo State (30 km drive from IITA Headquarters) on Day 2 of the training workshop (21 May 2015). The objectives of the field visit were to pre-test the survey instrument, allow the enumerators to practice everything that they had learned at the workshop, and undergo testing on the use of GPS and methods of DNA sample collection. The practical exercises during the field visit provided an opportunity for IITA to assess training outcomes. Furthermore, a post-field assessment session also gave an opportunity to the enumerators to receive feedback on their individual and group performances. The best of the enumerators were allowed to participate in the final survey.

CMS data collection and Collaboration with partners

Data collection began in June 2015 and ended in September 2015. Surveys were enumerated at two levels: Household and village. Implementation of all project activities involved the engagement and collaboration with NARs in Nigeria. IITA collaborated with the National Root Crops Research Institute (NRCRI), Umudike, on the development of the survey instruments, selection and training of enumerators, field data collection, and project management. IITA collaborated extensively with ADPs in all 16 states on field data collection and the management of social and cultural relationships with the study communities. Furthermore, IITA collaborated with Cornell University's Genome Diversity Facility on the analysis of DNA samples collected from the fields of interviewed farmers. We further undertook a qualitative case study in selected locations on gender differences and gendered decision-making in collaboration with international consultants to capture the role of gender in adoption and get a deeper understanding of adoption pathways, varietal preferences, and traits. Other end users that were involved in this segment of the study include a large number of commercial processors (medium-to-large scale or fully mechanized processors of garri, flour, instant fufu, native starch, ethanol, and feeds) whose preferences differ from those of households.

Picture 2. Survey enumerators carried out community social mobilization before starting interviews.

Project outcomes

As mentioned in the introduction and the Executive summary, the project has the following four major components: Part I: Large-scale household survey (2500 households from the states responsible for 80% of cassava production in Nigeria), II: DNA-based varietal identification using SNPs, III: GPS-assisted area measurements, and IV: Gender-differentiated end-user surveys on varietal and trait preferences. In the following section, we document the results from each component of the project.

Component 1: Large-scale household survey

Take-home message:

- On average, about 60% of farmers have adopted improved cassava varieties. Adoption rates across the different states of the country suggest a large spatial heterogeneity. In particular, adoption rate reaches as high as 79% in the Southwest region but was only about 31% in the Southeast.
- In terms of preference heterogeneity, more than 70% of the farmers considered varietal traits such as quality of garri or fufu/akpu, high root yield, big root size, high market demand, and early maturity traits as very important. In addition, a substantial share of respondents considered traits such as storage, white root color, and pest and disease resistance as very important. However, more than 40% of the farmers reported varietal traits such as yellow root color, root shape, starch content, and good tapioca taste as not important.
- The main determinants of adoption of improved cassava varieties can be categorized into the following groups: (i) Information market externalities: Both extension access and mobile ownership affect the probability of adopting improved cassava varieties positively. (ii) Credit market externalities: Both access to formal credit sources and membership in social networks that relax credit market constraints such as membership in informal saving and credit associations as well as membership in cooperatives affect the probability of adopting improved cassava varieties positively. (iii) Labor market externalities: Households with large household size are more likely to adopt improved cassava varieties. (iv) Heterogeneity in plot management: We found that soil fertility and plot management are important determinants of adoption. In particular, farmers are less likely to grow cassava on plots of medium fertility compared with highly fertile plots. Moreover, farmers are more likely to adopt improved varieties on plots managed by men or jointly by men and women than on plots managed by women only. (v) Preference heterogeneity: farmers who perceive traits such as quality of garri, root yield, and early maturity as highly important are more likely to adopt improved varieties. Surprisingly, we found a significant and negative effect on adoption for the trait starch content.
- The major determinants of disadoption include the following: distance to market, lack of planting material, availability of better improved cassava varieties, and pest and disease problems.

In Nigeria, achieving food security requires new and productive investments, innovations, and policy actions in agriculture since it is the predominant sector on which the majority of food insecure households directly depend for their livelihood (Koundouri et al. 2006; Alene et al. 2007; Alene and Coulibaly 2009; Alene 2010; Dercon and Christiaensen 2011; Suri 2011; Alene et al. 2012). One particular intervention in this regard is the widespread dissemination of new varieties as cassava is the most important food crop in the country (Abdoulaye et al. 2014). However, despite considerable efforts to improve productivity through the sustained promotion of new and improved cassava varieties, adoption rates have not been well documented (Abdoulaye et al. 2014). Moreover, there is scant empirical work and literature on the main determinants of adoption. One source of problems in identifying the determinants of adoption in the context of Nigeria is the lack of data at varietal level. Given the large gaps in the existing adoption databases for crop genetic improvement and policy-oriented research, the CMS project collected data at varietal level which will be important for updating the adoption database of improved varieties and for unpacking the determinants of adoption and disadoption. The International Institute of Tropical Agriculture (IITA) has invested substantially in generating improved cassava varieties in the past, but the extent of diffusion of such improved cassava technologies at varietal level has hardly been documented (Abdoulaye et al. 2014; Alene et al. 2012).

Picture 3. A survey enumerator conducting village level Focus Group Discussion with farmers in the Southeast.

These diffusion data on new cassava varieties are particularly important, since there is at present hardly any nationally representative data to document adoption and diffusion levels. Unpacking the determinants of adoption and disadoption is also crucial to identify options and to design “best fit interventions” instead of “one size fits all” interventions in the cassava sector. However, it must be noted that the outcome of this report is not an end in itself. Instead it is an instrument that answers questions relevant to policy using empirical data to support decision-makers in allocating resources in the most efficient way. This is particularly important since mapping possible demand-driven intervention areas is an integral part of planning and development for any development-oriented research institution such as IITA. With the focus on cassava, in this section we address the following empirical questions. What cassava varieties are households using? Why are some farmers not adopting improved cassava varieties? Why do farmers disadopt improved cassava varieties? Understanding how and why households decide what varieties to adopt is crucial for designing appropriate dissemination policies and for ex-post impact assessment of adoption.

Adoption of improved cassava varieties

Table 2 reports adoption rates of improved cassava varieties at country and regional levels. On average, about 60% of farmers have adopted improved cassava varieties. The distribution of adoption by gender reveals that the adoption rate among MHHs is about 61.5% while the adoption rate among FHHs is relatively low (about 48.6%). Adoption rates also show a large spatial heterogeneity. In particular, adoption rates reach as high as 79% in the Southwest region of the country while it is only about 31% in the Southeast region of the country.

Table 2. Adoption rate of improved cassava varieties in 2015 (%).

	Adoption rate (%)
Full sample (Nigeria)	59.5
North	59
Southwest	79.2
Southeast	30.7
South-South	69.4

General preferred traits of cassava

In CMS, we also collected detailed data on varietal trait preference as it may directly affect the adoption decision. Individual preference heterogeneity in cassava traits affects the farmer's adoption decision through various channels in addition to profit (Useche et al. 2009; Useche et al. 2013; Bajari and Benkard 2005; Hensher et al. 1999; Ladd & Suvannunt 1976). Especially in the context of Nigeria, genetic improvements on cassava can affect the taste of edible cassava varieties, which can then affect farmer utility and thus adoption through the consumption side. The purpose of capturing heterogeneity in varietal trait preferences is twofold: First, it helps to develop targeting policy interventions for addressing supply-side constraints (for example, credit, infrastructure, or market failure).

Secondly, it helps breeders to identify varieties which fit circumstances instead of advocating a "one size fits all" intervention (demand-driven breeding). Table 3 presents data on preference heterogeneity. This section asked households for their general trait preference for cassava. Preferences were not variety specific. In particular, the survey instrument asked each respondent to rank on the Likert scale the specific traits of cassava varieties from not important to very important for their adoption decision. Our results revealed that more than 70% of the farmers considered varietal traits such as good quality of garri and fufu/akpu, high root yield, big root size, high market demand, and early maturity to be very important. In addition, a substantial share of respondents considered traits such as storage, white root color, and resistance to pests and diseases as very important. However, more than 40% reported varietal traits such as yellow root color, root shape, and good tapioca taste as not important. It must be noted that these are indicators of correlation between adoption decisions and trait preferences. Any causal inference for trait preference should account for other confounding factors that also affect the adoption decision.

Table 3. Data on cassava varietal trait preference (%)

	Not important	Moderately important	Important	Very important
Makes good quality garri	3.3	2.3	13.8	80.6
High root yield	0.2	2.1	18.2	79.5
Big root size	0.4	3.4	18.1	78.1
Make good quality Fufu	4.7	3.8	15.8	75.7
High market demand	2.2	5.6	21.1	71.1
Stores well underground	3.79	5.12	23.46	67.63
White color roots	7.50	12.86	32.49	47.16
Pest and disease resistance	5.64	13.61	33.67	47.1
Early maturity	1	4.59	21.48	72.91
Yellow color	49.3	18.65	18.24	13.81
Poundability	16.6	11.44	30.06	41.9
Root shape (ease of loading)	31.16	29.06	25.55	14.23
High starch content	19.15	22.33	32.21	25.31
Good tapioca taste	43.61	9.15	19.55	27.69

Picture 4. Interviewing a farmer using Computer Assisted Personal Interview tools in the South-South region of Nigeria.

Overall, the data on trait preferences show that breeding programs need to account for production traits (such as early maturity, pest and disease resistance, and high root yield), consumption traits (such as quality of garri and fufu*),* market and other related preferences (such as white root color and high market demand). However from a breeding perspective, some inherent trade-offs may exist among the available options. As such, it is very important not only to discuss correlates in terms of trait preference but also to unpack determinants of adoption while controlling for preference heterogeneity. Such analysis helps in understanding the current most binding problems of adoption and ultimately contributes to the improvement of productivity and sustainability of the existing cassava production system.

Agronomic practices, access to information and planting material

CMS collected data on basic agronomic practices such as application of fertilizer, crop and plot management, and soil fertility status. In addition, data on planting material, membership in social networks, and access to credit and extension were collected. In terms of inorganic fertilizer application, our analysis revealed that about 36% of farmers apply fertilizer to cassava (Table 4). Moreover, we found a very high spatial heterogeneity as application rate is the lowest in the Southwest (about 12%) and highest in the Southeast (about 53.9%).

Another agronomic related practice that we considered is plot management as it directly affects the productivity level. Our survey results revealed that in about 50% of the farmer households, both husband and wife jointly managed cassava plots. There were fewer households (16%) where women only are

Table 4. Fertilizer application, plot management and soil fertility status (%).

		Cassava plot management			Soil fertility status		
	Application of fertilizer (%)	Men	Women	Both	Poor	Medium	Good
Full sample	36	35.3	15.9	48.9	2.3	23.8	73.9
North	42	41.8	4.6	53.7	0.4	16.9	82.7
Southwest	12	63.8	8.1	28	1.3	9.1	89.6
Southeast	48	23.8	25.8	50.4	3.1	36.5	60.4
South-South	37	20	20	60	3.7	26.5	69.8

responsible for the management of cassava plots. Like fertilizer, there seems to be a lot of spatial heterogeneity in terms of cassava plot management. For instance, in the North and Southwestern region, men tend to be prominent cassava plot managers while women tend to be prevalent in the Southeast and South-South regions. In terms of the soil fertility level of cassava plots, about 74% of farmers reported that the fertility level of the soil in their cassava plots is good. Note that these values are self-reported and hence there could be potential biases as what is good or poor soil is subject to the education level and experience of farmers. In fact, most farmers perceive that their cassava plots have either good or medium level of soil fertility with very few farmers reporting poor soil fertility status for cassava plots.

Access to planting material and duration analysis

Access to improved planting materials

Access to planting materials is an important issue in policy as it is the most important entry point for promoting improved cassava varieties. In this regard the CMS project collected data on access to planting materials from both formal and informal sources. What is really striking is that more than 70% of farmers reported that their primary source for improved planting material was social networks (friends, relatives, and neighbors). Developing a formal seed system is therefore crucial since access to planting material is a requisite for the adoption of productive and yield-enhancing varieties. Table 5 presents the main sources of planting material at the time of first adoption. Several findings are of note: overall, more than 70% of the farmers reported that they obtained planting material for improved cassava varieties from social networks. This shows that social norms, like the norm of reciprocity, play a prominent role in the distribution of planting material. This highlights the importance of social networks in contexts where farmers face limited access to credit and formal seed markets. Other important sources included extension and government sources (13%). Few farmers reported that they obtained planting material through nongovernmental organizations, processors, research institutes, the cassava market, and farmers' associations. Each of these sources accounted for less than 6% as a source of improved planting material. Disaggregated results over the different regions of the country suggest the same trend where planting materials are being distributed by informal social networks. Our descriptive analysis shows that informal sources are the locus of planting material for improved varieties. The contribution of local markets is rather insignificant. In addition, the role of private processor networks seems as insignificant as that of local markets in terms of the distribution of planting material. Generally, given the importance of cassava as a food crop and its prospect as an industrial crop, there is a great avenue for private seed dealers. Table 6 shows the dynamics of adoption of improved cassava varieties in Nigeria.

Table 5. Source of planting material for improved cassava varieties.

	Full sample	North	Southwest	Southeast	South-South
Family/Friends/Relatives/ Farmers/Neighbors	70.4	67.8	79.8	63.1	66
Extension/Government	12.6	13.2	8.1	14.8	16.0
Cassava market	5.7	6.3	2.5	12.8	5.2
Research institutes	4.6	5.1	4.8	3.0	4.7
NGO	3.7	3.5	2.5	4.4	4.7
Processors	2.4	3.3	2.1	1.5	2.3
Farmers' associations	0.3	0.3	0.2	0.5	0.5
Others	0.3	0.5	0.0	0.0	0.7

Table 6. Year of first planting an improved variety (% of farmers).

Years	Full sample	North	Southwest	Southeast	South-South
1951-1960	0.1	0.0	0.1	0.3	0.0
1961-1970	0.3	0.0	0.2	0.6	0.4
1971-1980	1.0	1.1	1.0	1.2	0.7
1981-1990	4.1	4.4	6.1	1.5	2.5
1991-2000	16.2	13.7	20.7	15.7	13.6
2001-2010	41.7	46.8	41.1	42.4	37.5
2011 and above	36.7	33.6	30.5	38.7	44.9

Picture 5. A van carrying cassava seed from one village to another in the South West.

Since the whole welfare and productivity effects of improved cassava varieties depend on the successful diffusion of improved cassava varietal adoption, examining varietal turn-over at the farm level is important. Our analysis shows that more than 78% of the farmers have grown improved cassava varieties at least once in the past. The proportion of farmers that have planted improved varieties in the past is highest in the Southeast (96%) region and lowest in the northern (71%) part of the country. The high level of exposure to improved cassava varieties suggests that examining why farmers replace one improved cassava variety with another is important. Looking into the dynamics of adoption, we also found that a large proportion of the farmers (58%) first planted improved cassava varieties between 1991 and 2010. Fewer farmers had access to improved varieties prior to 1970. Furthermore, the proportion of farmers that first planted improved varieties from 2011 was very impressive (37%).

However, the replenishment rate of planting material seems to be very low. Our survey results reveal that only 6% of adopters managed to replenish their planting material and about 94.1% have never done so. This result is directly linked to our previous finding about the lack of access to planting material from formal sources.

Table 7. Planting material replenishment rate.

	Full sample	North	Southwest	Southeast	South-South
Never renewed (%)	94.08	94.81	94.54	88.4	95.3

Table 8. Contact with extension agents (%).

	Full sample (%)	Female members (%)	Male members (%)
Contact with extension agents	29	18.7	32.5
Number of extension visits	5.03	1.57	3.47

Extension and credit access

As part of the CMS project, data on institutional and market-related aspects of cassava production were collected. Among these, data on access to extension and credit are reported following. Extension access is important for technology adoption as it reduces supply-side constraints that arise owing to information market inefficiencies. In particular, extension access facilitates the diffusion of improved cassava varieties by exposing farmers to the best farming and management practices (Wossen et al. 2013). For example, households who do not have access to private sources of information such as mobile phones and radio may learn about new improved cassava varieties from extension workers. Yet, extension access may also hinder the diffusion of improved cassava varieties as extension workers may exclude the poorest farmers or lack both the incentive and accountability needed for reliable information supply to meet the needs of smallholder farmers (Davis 2008). Our data show that farmers' access to extension services is very poor in Nigeria. In our data, only 29% of farmers have had at least one contact with extension services (Table 8).

The proportion of female members of the household that had contact with extension services was remarkably lower than that of the male members of the households. In fact only 18.7% of female members have had extension contact while the corresponding rate for male members is about 32.5%. This result suggests that extension advisory service is indeed gender-biased. Not only is access to extension service low but the frequency of extension visits is also very low. On average, the frequency of extension visits stands at about five days/year of which about 3.47 days are for male members and 1.57 days for female members. While examining the purpose of extension visits, we found that among those that had contacts with extension agents, 85% obtained training or advice about farm work and about 55% reported having received training and advice on cassava production. Although the differences between male and female members in terms of extension access and visits were significant, training of members by extension systems seems the same for both.

Picture 6. Survey enumerators battled odds to reach some remote villages that could not be reached by automobile.

Table 10 reports data on credit access. Data on access to credit are important as this is widely regarded as a requisite institutional innovation that can help to overcome the liquidity constraints that impede smallholders. According to our survey, about 40% of the farmers have access to credit. There was no significant difference in the level of credit access among male and female members of the households as about 38.2% of males and 32.5% of females have access to credit. For households with credit access, the principal sources seem to be savings and credit organizations/associations, relatives and non-relatives, and friends (Table 10). Both combined sources accounted for about 75% of credit facilities that were available to respondents.

In addition, about 10% of farmers reported having obtained credit from commercial banks. Surprisingly, the role of moneylenders as the primary source of credit is quite insignificant (about 5.6%). However, their role as a secondary source of credit is high, suggesting that farmers revert to traditional moneylenders as a last resort.

Our data further show that credit access related to cassava production seems to be very low. Only about 24% of the respondents reported that they obtained credit for cassava production. Credit access for key cassava inputs is even lower with only 18% (Table 11).

Table 9. Extension service on farm work and cassava production.

Type of advices	Full sample (%)	Female members (%)	Male members (%)
Whether Extension contact gave training or advice on cassava production	55.3	39.6	52.7
Whether Extension contact gave training or advice on farm work	85.1	88.8	83.7

Table 10. Sources of credit.

	Full sample (%)		Female members (%)		Male members (%)	
Sources	First source	Second source	First source	Second source	First source	Second source
Government	2.0	1.0	1.4	0.7	2.0	0.7
NGO	1.3	0.7	1.0	0.1	1.0	0.4
Savings and credit organizations/associations	42.8	23.9	41.0	24.6	40.8	23.9
Commercial banks	9.9	7.4	5.8	5.1	9.6	7.6
Friends, relatives and non-relatives	33.7	37.2	41.3	35.9	35.8	35.8
Money lender	5.6	10.4	6.1	10.4	5.9	11.8
Church/Mosque	1.5	4.0	1.7	5.2	1.6	3.8
Cassava Trader	0.6	0.8	0.2	0.7	0.2	0.6
Processor		0.2		0.6		0.1
Others	2.6	14.4	1.5	16.7	3.1	15.3

Table 11. Credit access for cassava related activities.

	Full sample	North	Southwest	Southeast	South-South
Ever obtained micro-credit for cassava production (%)	23.6	19.5	16.4	36.6	19.2
Obtained micro-credit to buy inputs for cassava production 2014 cropping season (%)	18.3	14.5	11.8	29.9	15.0

Table 12. Community-based associations (%).

Associations	Membership (%)	Female members (%)	Male members (%)
Cassava growers association	20.1	13.1	17.9
Religious group	85.0	81.8	78.1
GES farmers group	16.7	10.5	14.6
Women in agriculture	6.2	6.06	2.2
Credit and savings group	32.6	28.6	26.6
Cooperative	25.0	17.9	22
Mutual Aid group	36.8	31.7	33.2
Others	4.3	2.6	3.72

Membership of social networks and associations

Membership of social networks helps members to realize economic benefits (Rogers 1995; Wossen et al. 2013; Wossen et al 2015; Wossen et al. 2016; Munshi 2004; Young 2009; Conley and Udry 2010; Maertens and Barrett 2013; Bandiera and Rasul 2006). This is especially important as access to credit from formal institutions is very low, hence, informal social networks and institutions are important to relax the liquidity problem and also the information asymmetry that farmers face. Social networks, by providing social capital, facilitate the acquisition of resources that are essential for the adoption of improved cassava varieties (Wossen et al. 2015). Our analysis shows that the major community-based associations with the highest number of members are the religious groups. About 85% of the respondents reported that they belonged to religious groups (Table 12).

In addition, a significant share of farmers also reported being members of cooperative societies and other important groups such as those for mutual aid and credit and savings. Membership in informal associations seems to be random among family members as there is an insignificant difference between male and female members.

Utilization of cassava

In Nigeria, cassava is the source of livelihood for farmers and countless processors and traders as a cash crop and source of livestock feed. Herein, we report the main uses of cassava based on CMS data. Results from CMS show that cassava is used mainly for food and as a cash crop. For instance, about 50% of the cassava produced in Nigeria is sold for cash income (Table 13). In the North and Southwest regions, about 60% of the farmers have reported selling cassava as a cash crop. In the southern part of the country, less than 50% of the farmers have reported selling cassava for additional cash income. In addition to cash income, significant numbers of farmers have also used cassava as food. Our survey result suggests that about 40% of the quantity produced was consumed as food. The data further show a large spatial heterogeneity in the utilization of cassava. For instance, about 45% of the farmers in the South-South and Southeast have utilized cassava as food but only 30% in the North and Southwest. Our result also suggests that less than 10% of the total quantity produced was used as a gift by the study participants.

Since the main use of cassava seems to be as a cash crop, we further examined the share derived from sales in the total household income. About 40% of farmers have reported that cassava contributed about half of their income (Table 14). In about one-third of the households, it accounts for about 75% of their total income. In about 19%, it contributed a quarter of the income. Only a small number reported that cassava contributed 100% of the income earned.

Furthermore, while considering the role of cassava among women, we found that in about one-third of the households, cassava contributed 50% of the total income of women. Again, in about a quarter of the households it contributed 25% of the income earned by the women. However, in about 11% of households, it was reported that cassava did not contribute anything to the income of women.

Table 13. Cassava utilization (%).

	Full sample	North	South- West	South- East	South- South
Percentage of cassava used for sales in the household	52.9	58.6	58.9	47.2	46.9
Percentage of cassava used for consumption in the household	38.0	32.2	31.0	44.4	44.6
Percentage of cassava used in other ways in the household such as gifts	9.1	9.3	10.1	8.4	8.6

Table 14. Share of cassava income from total household income.

Share (%)	Full sample	North	Southwest	Southeast	South-South
100	4.5	2.9	4.7	5.3	5.1
75	31.9	30.9	28.6	34.5	33.6
50	38.0	37.0	44.1	34.9	35.8
25	18.7	22.9	18.1	15.7	17.8
10	3.9	5.9	1.9	3.2	4.7
0	3.1	0.3	2.6	6.4	2.9

Table 15. Share of cassava from total household food consumption.

Share (%)	Full sample	North	Southwest	Southeast	South-South
100	1.2	0.5	1.8	1.9	0.8
75	32.3	18.9	21.3	34.1	55.3
50	34.9	27.6	31.3	43.4	37.3
25	24.3	38.1	33.7	19.0	5.9
10	7.1	14.4	11.6	1.6	0.7
0	0.2	0.5	0.3	0.0	0.0

Picture 7. Local cassava processing into lafun.

These are most likely in the households where the women members are engaged in non-farm activities. While examining the use of cassava as food, we found that a significant share of household consumption is derived from cassava. CMS data show that in about one-third of the households, cassava contributes 75% of the food consumed by family members (Table 15). Similarly, another one-third of respondents reported that the crop contributed about 50% of the food consumed by household members. More than 90% of households reported that cassava contributed at least 25% of the food consumed. The number that reported that cassava did not contribute to household feeding was grossly insignificant (only 0.2%).

Disadoption among farmers

In mainstream agricultural economics, technology adoption has been a central research issue. However, research on the issue of disadoption has been rather scanty. Observation in many developing countries suggests that disadoption occurs for many types of agricultural technologies. Disadoption may occur owing to the life cycle of the technology, supply-side constraints, or a reduction in profitability. As part of the CMS, we asked respondents to list all improved varieties that they had tried in the past. Similarly, data on the list of varieties that farmers no longer grow were collected. Other reasons for disadoption include varieties not being good for processing (6%) and having a low market price (5%). The farmers also mentioned the following as some of the factors that made them disadopt improved varieties: susceptibility to pests or diseases, lack of money to buy stems, requiring too much labor/work to grow, late maturity, short underground storage period after maturity, and poor mealy qualities/poor starch content (the lafun made from the varieties turns to water very quickly). Furthermore, persistent destruction of farms by Fulani cattle, especially in the North, is also mentioned as a reason for disadoption.

Variety-specific trait preference heterogeneity

Studies that consider technology adoption and disadoption largely ignore the role that varietal traits play in the adoption decision. As mentioned at the beginning of this section, given the role of preference heterogeneity for the adoption of improved cassava varieties, the CMS project collected data on the main consumption, production, and processing traits of existing improved cassava varieties. Herein, we report the results for the production, consumption, and processing traits of cassava varieties currently grown by farmers. The analysis in this section differs from that of the section where we examined general preferences as it considers the desirable traits of improved cassava varieties that are grown by farmers in the current production year.

Production traits

The survey instrument asked households to report the first, second and third most important production traits that farmers like for the varieties they grow on their plot. Our results suggest that early maturity, high yield, and tuber size are the major preferred production traits (Table 17). Others include resistance to drought, long storage underground after maturity, resistance to diseases, stem height, and white roots.

Processing traits

In addition to production traits, the survey instrument also asked households to report the first, second, and third most important preferred processing traits for the varieties grown by farmers. Our descriptive results (Table 18) revealed that the most preferred processing traits are the ability of varieties to be processed into garri or fufu/akpu and ease of peeling.

Consumption traits

The major preferred consumption traits that emerged from our descriptive analysis are taste of varieties for garri, fufu, good poundability, and palatability when boiled (Table 19). Other traits that were considered to be important by the respondents include smoothness, fiber content and white color.

Table 16. Reasons for disadoption (%).

Reasons	Pooled	North	Southwest	Southeast	South-South
Stems not available	25.4	30.3	22.7	24.0	25.5
Prefer other variety	18.3	21.2	20.0	16.0	11.8
Low yielding	15.5	9.1	16.4	16.0	21.6
Not good for processing (poor starch/food quality)	6.3	6.1	5.5	8.0	7.8
Low market price	4.8	1.5	6.4	0.0	7.8
Poor adaptation to soil and agro-ecology	4.8	3.0	4.5	8.0	5.9
Poor underground storage after maturation	4.4	3.0	5.5	0.0	5.9
Late maturity	4.0	1.5	4.5	8.0	3.9
Susceptible to pests and diseases	3.2	3.0	1.8	8.0	3.9
Lack of cash to buy stems	2.8	0.0	4.5	8.0	0.0
Require too much work	1.6	1.5	1.8	4.0	0.0
Poor taste	0.4	1.5	0.0	0.0	0.0
Don’t know	1.2	1.5	0.9	0.0	2.0
Others	7.6	16.7	5.5	0.0	4.0

Table 17. Most preferred production traits.

	What is the first production trait you like MOST about this variety?	What is the second production trait you like MOST about this variety?	What is the third production trait you like MOST about this variety?
Resistance to drought	5.85	2.38	4.33
Resistance to pest	1.51	1.56	2.38
Resistance to disease	4.01	3.5	5.49
Early maturity	37.49	17.23	12.71
High yielding (roots)	21.93	27.18	16.75
Stem height	3.64	5.8	6.71
Root size	16.34	27.26	23.47
Stores well underground	2.98	5.84	10.69
White cassava roots	2.13	4.69	8.74
Erect type	0.44	0.29	0.65
Yellow cassava roots	0.44	0.78	1.01
High yield stem	0.85	1.40	3.32
Spread/branched type	1.99	1.60	3.25
Other specify	0.40	0.49	0.51

Table 18. Preferred processing traits by household heads (%).

	What is the first processing trait you like MOST about this variety?	What is the second processing trait you like MOST about this variety?	What is the third processing trait you like MOST about this variety?
Ease of peeling	20.02	4.62	12.84
Low water content	8.75	7.18	5.58
Ability to be processed into garri	42.42	23.39	9.81
Ability to be processed into fufu	10.31	38.94	21.13
Swells well	3.63	4.74	6.54
Molds well	2.67	2.81	5.26
The color is fine	5.01	5.86	10.05
It is heavy for garri	2	2.89	5.90
Draws well for fufu	1.41	2.85	4.78
Yellow roots	0.48	0.41	0.64
White roots	2.22	4.04	7.18
Cream roots	0.11	0.12	0.16
Ability to be processed into tapioca	0.56	1.61	9.49
Others	0.41	0.54	0.64

Other processing traits that are deemed to be important include low water content, fine color, and white roots.

Table 19. Preferred consumption traits by household heads (%).

	What is the first consumption trait you like MOST about this variety?	What is the second consumption trait you like MOST about this variety?	What is the third consumption trait you like MOST about this variety?
Good pounding ability	9.36	3.98	7.89
Fiber content	4.74	2.03	3.36
Taste for garri	60.16	19.2	7.02
Taste for fufu	11.21	47.58	13.96
Palatability when boiled	4.59	5.16	10.23
Cooking time	1.78	3.09	4.31
Smoothness	2.33	4.14	8.04
Palatability when roasted	0.22	0.81	1.68
White color	1.96	4.47	10.96
Yellow color	0.59	0.77	0.73
Taste of tapioca	0.78	2.96	12.65
Taste of lafun	1.59	5.03	17.84
Others	0.7	0.77	1.32

Determinants of adoption and disadoption: econometric analysis

In the previous section, we present the relevant aspects of household and variety specific characteristics. However, inference about the main determinants of adoption and disadoption cannot be done by simply comparing adopters and non-adopters. As such, we used standard micro-econometrics techniques to unpack the main determinants of adoption and as well as the determinants of intensity of adoption (Feder et al. 1985). In particular, we employed a double-hurdle model that allows the incorporation of the probability of adoption and the intensity of adoption in a separate process through probit/logit models in the first step and truncated normal model in the second step (Croppenstedt et al. 2003; Griliches 1957; Wossen et al. 2015). This procedure has a considerable advantage over separate probt/tobit models as it allows the direction and significance of parameter estimates for a given variable of interest to be different in the first and second stages. The double-hurdle model is specified in the following functional form:

$$v, y|x_1, x_2) = \{1 - \varphi(x_1\gamma)\}^{1(w=0)} \left[\varphi(x_1\gamma)(2\pi)^{-\frac{1}{2\sigma^{-1}}} \exp\{-(y - x_2\theta)^2 /2\sigma^2\} /\varphi(\theta x_2/\sigma\right]^{1(w=1)} \quad (1).$$

This equation incorporates a probit model to determine the probability of adoption (y) > 0 and a truncated model for a given positive value of y (intensity of adoption). In this model is a binary variable which takes a value of one for adopters and zero otherwise. The variables x_1, x_2 include the determinants of adoption and intensity of adoption. These include access to information and credit, membership in social networks, plot- and household-specific characteristics, and stated preferences for improved cassava varieties. Table 20 presents the descriptive statistics of the key variables of interest. Based on the economic theory, we included control variables at the household and village level. We included household characteristics such as age, household size, education, membership of different social groups and wealth indicators such as TLU[2] and land size. In addition, household-specific improved cassava variety preferences are included as additional controls. The descriptive statistics shows that the average household size is about 4.5 members and the average household head is 51 years old. Moreover, most of the respondents are literate, with an average of nine years of schooling for the household-head. About 89% of the respondents are married. Smallholder farmers are located on average 3 km away from a nearby village market. In terms of membership of local social networks, the data shows that about 25% and 34% of farm households are members of cooperative and informal credit and saving associations, respectively. Similarly, about 40% of farmers are members of mutual aids.

[2] Livestock Unit.

In terms of institutional variables, about 65% of farmers have access to credit facilities and 50%[3] to extension services. Note that, since access to credit is potentially an endogenous variable, we used a household level question on credit access to construct a village level credit access indicator which is arguably exogenous to household characteristics. We used the indicator to distinguish villages with relatively easy access to credit from those without access. We set a threshold of 25% to distinguish "credit villages" from "non-credit villages". At a threshold of 25%, a village is considered to be a "credit village" if more than 25% of the households respond positively to having access to credit. As such, the credit access value (65%) reported earlier is based on the 25% threshold. Similarly, to control for selection bias caused by observable and unobservable factors in extension services, we aggregated it at the village level, with the presumption that village-level extension access would be exogenous to individual household characteristics. We constructed an indicator to distinguish villages with relatively easy access to extension services from those without access.

Table 20. Descriptive statistics.

Variable	Full sample	Adopters	Non-adopters	Mean diff
Household size	4.54	4.8	4.3	0.47***
Education (years of education)	8.75	9.03	8.43	0.594***
Age	51.5	50.9	52	-1.09***
Sex (1=male)	0.89	0.91	0.87	0.037***
TLU	0.7	0.85	0.52	0.33
Access to credit (1=yes)	0.65	0.71	0.58	0.13***
Land size (ha)	5.6	6.1	5.07	1.01**
Access to extension (1=yes)	0.5	0.61	0.36	0.25***
Distance from market (km)	2.87	3.01	2.69	0.32**
Availability of cassava stems (1=yes)	0.51	0.44	0.59	-0.15***
Radio ownership (1= yes)	0.84	0.85	0.82	0.03***
Mobile phone ownership (1= yes)	0.97	0.98	0.95	0.027***
TV set ownership (1= yes)	0.732	0.731	0.734	-0.0032
Membership in credit and savings groups (1=yes)	0.34	0.36	0.31	0.05***
Membership in cooperatives	0.25	0.30	0.19	0.11***
Membership in mutual aid groups	0.40	0.402	0.398	0.004
Poor soil	0.02	0.0198	0.023	-0.003
Medium soil	0.23	0.19	0.28	0.092***
Good soil	0.75	0.79	0.69	0.1***
Men managed plots	0.37	0.43	0.30	0.13***
Women managed plots	0.15	0.11	0.20	0.09***
Men and women managed plots	0.48	0.46	0.50	-0.043***
Poundability	3.05	2.93	3.19	0.253***
Quality of garri	3.7	3.8	3.6	0.2***
Starch content	2.65	2.58	2.7	-0.12***
Quality of Fufu	3.68	3.69	3.67	0.014
Tapioca taste	2.42	2.28	2.6	-0.32***
Root yield	3.7	3.8	3.65	0.15***
Early maturity	3.65	3.72	3.56	0.16***
Easy to load	2.24	2.25	2.22	0.03
Market demand	3.6	3.65	3.53	0.12***
White color	3.22	3.18	3.29	-0.11***
Stores well	3.56	3.55	3.56	0.01
Pest and disease resistance	3.16	3.2	3.1	0.1***
Stem yield	3.03	2.99	3.09	0.1***
Yellow root color	1.95	1.97	1.92	0.05

****, ** and * refer to significant at 1%, 5%, and 10% level.*

3 Note that only 29% of households have access to extension, when it is measured at the household level.

Moreover, on average, adopters own more land, have better credit and market access, a larger family size, and are more educated. Similarly, in terms of membership of social networks, 30% of cooperative members have adopted improved cassava varieties while only 19% of non-members are adopters. This difference is also statistically significant at 1%. In addition, adopters were also significantly different from non-adopters in terms of membership in mutual aid groups and informal credit and savings associations. Adopters and non-adopters were also significantly different in other characteristics such as plot management and soil type.

In addition to socioeconomic characteristics, we also included important household level preference heterogeneity for important cassava traits. These include production, consumption and processing traits measured on a Likert scale of 1 to 4[4]. In particular, the survey asked both adopters and non-adopters for the main production, consumption and processing traits on a 4-point scale from very important to not important. Table 20 shows the average scores with regard to different production, consumption, and processing attributes as scored by farmers. The traits we included in the regression approach are production traits (starch content, root and stem yield, maturity date, pest and disease resistance), consumption traits (quality of fufu*,* and garri and tapioca taste), processing traits (poundability, ease of loading, etc.) In addition, Table 20 presents mean differences in terms of preference heterogeneity between adopters and non-adopters. We found significant differences among them in the perception of production traits such as root and stem yields, starch content, resistance to pests and diseases, and early maturity. This suggests that adopters and non-adopters have heterogeneous preferences for production traits.

Results on the determinants of adoption

Table 21, presents the result of our regression analysis in which we estimated the determinants of improved cassava adoption as well as the intensity of adoption at the plot level. Among the socioeconomic factors, household size and sex of the household head have a positive and statistically significant effect on the probability of adoption. However, these variables play no significant role for intensification. Rather, older household heads tend to allocate more land to improved cassava than younger household heads. This effect is non-linear and decreases with age. Access to extension has a positive and statistically significant effect on the probability and intensity of adoption. The positive and significant effect of extension shows the role that proactive extension can play in accelerating the adoption and intensification of improved cassava varieties. Similarly, the effect of credit access is positive and statistically significant albeit only for adoption but not intensification. We did not find any significant effect of credit access on the decision to intensify improved cassava varieties. This suggests that given the choice, relaxing information constraints through a targeted extension delivery system would be more effective than relaxing liquidity constraints. In terms of membership of informal social networks, we found positive and statistically significant effect for membership in informal saving and credit associations as well for membership of cooperatives on the probability of adopting improved cassava varieties.

Table 21. Determinants of adoption.

	Based on farmers self-report	
Variable	Coefficient	Z value
Household size	0.020**	2.1
Adoption (=1 if yes and 0, otherwise)		
Education	0.001	0.25
Age	-0.005	-0.5
Age2	0.000	0.77
Sex	0.164*	1.87
TLU	0.000	-0.17
Access to credit	0.146***	3.16
Access to extension	0.176***	3.55
Distance from market	-0.003	-0.72

[4] 1 =Very importnat, 2= Important, 3=Moderately important 4=Not important.

	Based on farmers self-report	
Variable	Coefficient	Z value
Availability of cassava stems	0.173***	3.45
Radio ownership	0.020	0.35
Mobile phone ownership	0.726***	5.55
Membership in credit and savings groups	0.146***	3.22
Membership in cooperatives	0.120**	2.44
Poor soil	0.103	0.72
Medium soil	-0.099**	-1.98
Men managed plots	0.143*	1.84
Men and women managed plots	0.225***	2.74
Poundability	-0.018	-0.78
Quality of garri	0.199***	5.67
Quality of fufu	-0.056	-1.6
Easy to load	-0.023	-1.07
Starch content	-0.067***	-3.1
Tapioca taste	0.086***	3.61
Market demand	-0.012	-0.35
Root size	0.066	1.46
White color	-0.015	-0.58
Stores well	0.022	0.74
Root yield	0.109***	2.21
Pest resistance	-0.016	-0.62
Stem yield	-0.067***	-2.76
Early maturity	0.098***	2.69
Yellow roots	0.012	0.58
Intensity of adoption		
Household Size	-0.002	-0.8
Education	-0.001	-0.96
Age	0.011***	3.51
age2	-0.0001***	-3.63
Sex	0.052*	1.66
TLU	0.000	0.16
Access to credit	0.00002	0.01
Access to extension	0.032**	2.03
Distance from market	0.003*	1.85
Availability of cassava stems	0.053***	3.28
Radio ownership	-0.016	-0.85
Mobile phone ownership	-0.082	-1.56
Membership in Credit and Savings groups	0.023	1.62
Membership in cooperatives	0.035**	2.31
Poor soil	-0.018	-0.39
Medium soil	-0.015	-0.88
Men managed plots	-0.070**	-2.51
Men and women managed plots	-0.024	-0.84
Poundability	0.000	-0.01
Quality of garri	0.022*	1.74
Quality of fufu	-0.001	-0.07
Easy to load	-0.010	-1.49
Starch content	0.012*	1.78
Tapioca taste	0.002	0.22
Market demand	0.007	0.66
Root size	-0.013	-0.9
White color	-0.005	-0.71
Stores well	-0.038***	-4.21
Root yield	0.001	0.04
Pest resistance	0.010	1.22
Stem yield	-0.018**	-2.37
Early maturity	0.011	0.95
Yellow root	-0.013**	-2.01
N	4468	

*Regional dummies are included but not reported. ***, **, and * refer to significant at 1%, 5% and 10% level.*

For the intensification decision, we again found that membership of cooperatives has a positive and significant effect. These results underscore the crucial role of social networks in enhancing adoption. In terms of access to information through mobile phones and radio, we did not find any statistically significant effect for radio ownership. However, farmers who own mobile phones are more likely to adopt improved cassava varieties than those who do not own mobile phones. The primary pathway through which mobile telephony affects adoption decisions is by relaxing the information market imperfections which are the intrinsic feature of agriculture in rural Nigeria. This is apparent as mobile phone ownership has no significant effect on the decision to intensify the use of improved varieties. This result underscores the fact that addressing information market imperfection in agriculture where there is a significant lag between the production decision and output realization is crucial for enhancing adoption. In addition, production of cassava, as any other crop, is inherently stochastic due to weather shocks, pest infestations, and other shocks, which affect the general market supply conditions and therefore adoption decisions. As such, mobile ownership helps farmers to gather and process information which is crucial in adopting improved varieties. Another important variable that affects the adoption and intensification decisions is access to cassava stem markets. Our results show that households who have access to cassava stem markets are more likely to adopt improved varieties. In addition, farmers who have access intensify their production compared with those without access.

In terms of plot-level variables, we found that soil fertility and plot management are important determinants of adoption. Our results suggest that farmers are less likely to grow cassava on plots with medium fertility compared with highly fertile plots. Note that, soil fertility is measured by the farmers' own self-perception and hence may be subject to measurement error as what is "fertile" can be perceived differently by different farmers. For plot management, we found that farmers are more likely to adopt improved varieties in plots managed by men or jointly by men and women than in the plots managed by women alone. Surprisingly, farmers intensify cassava production on women-managed plots compared with plots managed by men.

Coming to preference heterogeneity in cassava traits (production, consumption, and processing), we indeed find significant effects for most traits. As mentioned before, unlike previous studies that didn't explicitly capture the role of individual traits of the technology in farmers' adoption decisions, our regression explicitly uses preference heterogeneity as additional controls in the regression. This is particularly important as improved varieties have a wide range of traits (e.g., the potential to reduce pests and diseases, increase yields, improve quality of taste, early maturity, etc.). As such, farmers assess not only the resource constraints they face but also the benefits and costs associated with the distinct trait combinations of various improved varieties in their adoption decision. Our results in Table 21 show that individual preference heterogeneity in improved varieties affects the decision on adoption as well as on intensification. In particular, those farmers who value highly such traits as quality of garri, root yield, and early maturity are more likely to adopt improved varieties. Surprisingly, we find a significant and negative effect on adoption for the trait, "starch content". This might be due to the lack of local processors for starch. In terms of an intensification decision, only preference for garri and starch content become significant, suggesting that farmers with a higher preference for garri and starch content tend to intensify production of improved cassava.

Determinants of disadoption

In this section, we present probit results on the determinants of disadoption. In our data set we measured disadoption by a dummy variable which takes a value of one if the household stops growing a particular improved cassava variety during the current production year. In particular, the survey instrument asked respondents whether they had ever grown an improved variety and whether they were still growing it. Disadoption occurs when a household has grown the variety at least once until last year and stopped growing it in the current production year. We estimated the model at the household level as disadoption variable was measured at the household level. Accordingly, about 11.6% of the farmers have stopped growing a particular improved cassava variety in the last production year[5]. Farmers mentioned non-availability of planting materials, preference for other varieties, varieties not being good for processing and having a low market price as the main reasons for disadoption. Our probit results reveal that variables such as access to extension, distance to market, lack of planting material, availability of better improved cassava varieties and pest and disease problems are statistically significant.

The positive and statistically significant effect of extension access is surprising as we also found similar effects for adoption. This result suggests that extension plays a positive role for adoption and disadoption. Note that care has to be taken in interpreting this result as disadoption does not necessarily mean switching from improved cultivars to local varieties. It may be that those farmers with extension access disadopted some improved varieties and grew better varieties. In fact, one of the main reasons why farmers disadopted improved varieties is the availability of better varieties. As expected, the effect of distance to market is positive and significant. Farmers living far away from market are less likely to benefit from adoption due to higher transaction costs. Lack of planting material has also a positive and statistically significant effect on disadoption. In particular, a lack of planting material increases by 34% the probability of disadoption. Similarly, access to better improved varieties increases by 27% the probability of disadoption. Finally, the prevalence of pests and diseases increases the probability of disadoption by 28%.

Table 22. Determinants of disadoption.

Variable	Parameter estimates	Marginal effects
Household size	0.0255	0.0034
	(0.025)	(0.0034)
Education	(0.118)	-0.00069
	-0.0005	(0.0015)
Age	0.0010	0.00013
	(0.004)	(0.0005)
Sex	0.0845	0.0114
	(0.210)	(0.028)
TLU	-0.0014	-0.00018
	(0.002)	(0.0003)
Access to credit	-0.0442	-0.0059
	(0.114)	(0.015)
Access to extension	0.2810**	0.038**
	(0.118)	(0.016)
Distance from market	0.0118**	0.0016**
	(0.006)	(0.0008)
PPI score	0.0061	0.0008
	(0.005)	(0.0007)
Radio ownership	-0.1230	-0.017
	(0.150)	(0.022)
Mobile phone ownership	-0.2221	-0.03
	(0.247)	(0.033)
Membership to credit and saving groups	-0.0792	-0.011
	(0.111)	(0.015)
Membership to cooperatives	0.0008	0.0001
	(0.112)	(0.015)
Lack of planting material	2.5339***	0.343***
	(0.274)	(0.037)
Pest and disease problem	2.0997***	0.28***
	(0.471)	(0.063)
Availability of better variety	1.9815***	0.268***
	(0.246)	(0.032)
N	1566	1566

*Regional dummies included but not reported. Standard errors clustered at numeration area level are reported in parentheses, *** p<0.01, ** p<0.05 * p<0.1. Variables not reported.*

[5] Note that, stopping growing a particular improved variety may also mean that the households switch to a better improved variety.

Component 2: DNA-based varietal identification using SNPs

Introduction

Several options exist in varietal identification for tracking the adoption of improved varieties. However most of these methods have inherent uncertainty levels and estimates often have wide confidence intervals. The traditional variety identification methods include the following:

- Secondary sources (published government reports, available data sets).
- Seed sales and seed multiplication/distribution data.
- Expert opinion / key informant interviews.
- Community-level surveys (i.e., focus group discussions).
- Farmer elicitation: asking farmers what varieties they planted (name and/or type) and collecting this and supplemental information as part of a household survey.

The DNA fingerprinting techniques offer a reliable method to accurately identify varieties grown by farmers and increases accuracy and credibility in the interpretation of adoption rates and associated economic analysis. Unlike phenotype-based methods, DNA is independent of environmental conditions or plant growth stage and more abundant than morphological descriptors.

The DNA fingerprinting of all cassava cultivars collected from 2500 households was one of the main components of the CMS project. The main activities in this component included sample kit and booklet preparation, field sample collection, establishment of a tracking system from field to laboratory, implementing high throughput DNA extraction, genotyping, bioinformatics, and variety identification analysis. The report will focus on details of samples collected and the methodology followed for the analysis and output of the research, in particular for the identification of improved varieties growing on farmers' fields, and determination of their extent across different cassava-growing regions of Nigeria.

Picture 8. Survey enumerators collecting cassava leaf samples in the field.

An independent detailed manual in relation to large-scale sample collection, preservation, tracking, and DNA extraction has been developed.

Overview of the DNA fingerprinting workflow

The DNA fingerprinting component of the CMS involved establishment of a clear workflow which could be used as a reference for similar studies addressing the tracking of adoption of improved varieties. The workflow was as follows.:

1. Establishment of a reference library comprising improved varieties, the genebank collection, and landraces.
2. Field sample collection and preservation.
3. Establishment of tracking system to ensure the chain of custody of sample identification from field to laboratory.
4. High throughput DNA extraction that allows extraction daily from a large number of samples.
5. Genotyping.
6. Bioinformatics and cultivar identification.

Main activities in the DNA fingerprinting component

The DNA finger- printing component of CMS involved several activities from field sample collection to varietal identification and analysis. Following are detailed reports on the completed activities.

Sample and sample associated data collection

The study was conducted in 2500 sampled households across 16 states in four regions, North Central, Southwest, Southeast, and South-South, representing about 80% of the cassava production areas in Nigeria and consisting of 625 households per region. Leaf samples for all the varieties identified by farmers in each household were collected and preserved in plastic tubes containing silica gel and transported to the Bioscience laboratory at IITA in Ibadan, for DNA extraction. Household information including ID for region, state, local government area, EA, and household as well as the name of the household head was captured in a booklet. In addition, information on variety name, cropping pattern (mono-cropped or intercropped), field and plot identification, plot size in all the fields owned by the household was recorded and the GPS coordinates were measured of the household where the survey took place and the farmers' fields.

Establishment of sample tracking system

A standard tracking system is important particularly when dealing with a large sample size to reduce any possible introduction of human errors of sample mismatch and mix ups. A tracking system with multiple layers was implemented using barcode labels, self-adhesive stickers, booklet and tablet computers for capturing samples and sample-associated information. This process has improved the accuracy and reliability of the data. Duplicate barcodes were prepared and pasted both on sample collection tube and booklet for each sample collected. Once received in the lab samples were arranged in a set of

Picture 9. Recommended size and age of cassava leaf tissue, silica gel dried leaf, and sample identification.

96 on a plate made in-house and assigned distinct plate numbers. Both plate and sample numbers were written on the cover of each sample collection tube and were different for the 96 samples. Information on each vial consisting of the barcode label and other metadata was captured manually on a tablet and on a hard copy plate map. The barcode label on each tube was also captured with a barcode reader in parallel using DNA plate software obtained from www.wheatgenomics.org. Manually entered and barcode reader information was cross-checked. This has helped a lot in reducing the possible introduction of human errors or sample mix ups.

Summary on the samples collected

Field samples included leaf tissue collection, preservation in plastic tubes containing silica gel/desiccant and the recording of sample-associated information. Due consideration and intensive training were given to enumerators prior to the field visit and on the field to ensure proper sampling, conservation of plant tissue, and capture of sample related information. Before field visits more than 8000 plastic tubes of 50 ml size containing 20 g silica gel were prepared and adhesive labels with adequate space to capture information were pasted on all tubes. In addition a barcode label unique for each variety was prepared in duplicate and pasted on the sample collection tubes and booklets. Other sample-associated information was also captured. A total of 7428 different samples were collected from 2500 households of the four regions of Nigeria (Table 23). The largest number of samples was collected from the North Central followed by South-South, Southwest, and Southeast regions in descending order of sample size.

DNA extraction and genotyping by sequencing

DNA was isolated following the extraction protocol of Dellaporta *et al.* 1983) with some modification for the large-scale sample from a total of 7428 genotypes collected from 2500 households including quality control (89 clones genotyped in duplicate) and re-extracted samples. In-house modified protocol (Rabbi *et al*. 2014) was implemented that enables up to 10 plates of 96 samples each to be extracted per day. All the extracted DNA samples were quantified using the spectrophotometer and agarose gel electrophoresis for quality and quantity assessments. Test digestion with restriction enzyme was performed for 10% of the samples extracted as suggested by the Genetic Diversity Facility (GDF) at Cornell University for standard Genotyping by Sequencing (GBS) library preparation. DNA samples with high concentrations were diluted to 1000ng /μl. All extracted samples that passed the minimum quantity requirement (300ng /μl) were shipped to GDF for GBS. Samples were originally processed in three batches consisting of 30 plates for the first two batches and 20 plates for the third. A few of the total sequenced samples from Batch One were with a low pass count (< 500,000 reads) (Fig. 2). These samples were therefore re-extracted and included in the third batch. As a result a total of 7565 samples were genotyped including 89 samples genotyped in duplicate (Table 24). The purpose of genotyping samples in duplicate was to identify any possible SNP genotype error resulting from miscalling some heterozygous SNPs with low sequencing read depth as homozygotes and to determine a threshold to declare if two genotypes are the same or different (Table 24).

For GBS library preparation, the ApekI restriction enzyme (recognition site: G|CWCG) was used that produces less variable distributions of read depth and therefore a larger number of scorable SNPs in cassava. Eighty 96-plex GBS libraries were constructed following the standard procedure (Elshire *et al.* 2011) and sequenced at the Institute of Genomic Diversity at Cornell University using the Illumina HiSeq2500.

Table 23. Total number of samples collected from the respective region, enumeration area and household.

Region	Total number of enumeration areas	Total number of house hold	Total number of samples collected
North-Central	125	625	1996
Southwest	125	625	1775
Southeast	125	625	1740
South-South	125	625	1917
Total	500	2500	7428

Table 24. Summary on the number of samples genotyped by sequencing.

Sample source	Batch 1	Batch 2	Batch 3	Total number of samples extracted
North	497	931	568	1996
Southwest	758	921	96	1775
Southeast	1253	1	386	1740
South-South	342	967	608	1917
Quality control samples		30	59	89
Low pass count			48	48
Total	2850	2850	1765	7565

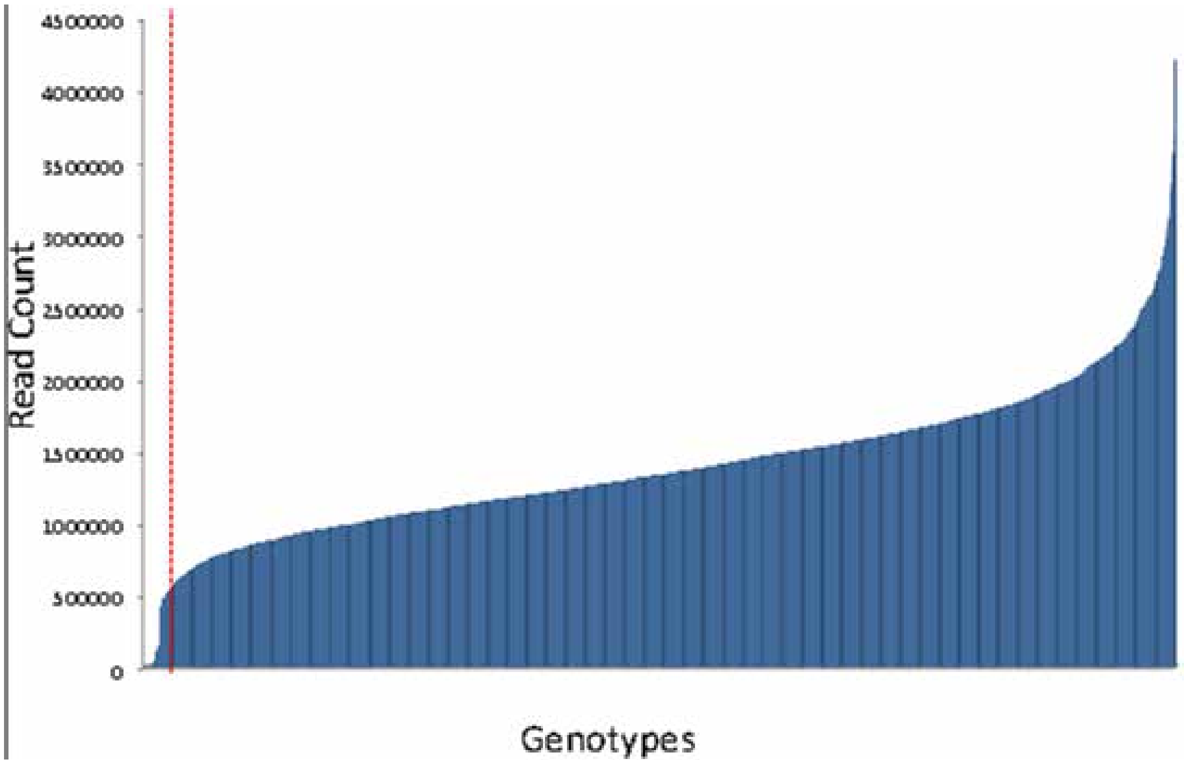


Figure 2. Number of sequence reads across genotyped samples. The dashed red line indicates the minimum pass count.

Development of a reference library for varietal identification

Developing a well curated comprehensive reference library in collaboration with the breeding programs is very important for tracking genotypes of interest, otherwise DNA fingerprinting alone can be used only to establish baseline data. The quality of the reference library (genotype traceability, and comprehensiveness) determines the "level of success" in varietal identification. The already available reference library with a total of 3891 diverse genotypes comprising a collection of known improved lines, IITA regional breeding program, IITA germplasm collection, wild species, and CIAT collection were used in the current study (Fig. 3). The GBS technique was applied for the development of the reference library. Likewise, the GBS procedure was implemented for the genotyping of the current samples collected from farmers for varietal identification. Raw sequences of the two data, consisting of accessions in the reference library and the current samples collected from farmers' fields, were combined for single nucleotide polymorphism (SNP) calling and further identification and analysis of varieties.

Picture 10. A research assistant demonstrates the procedure of leaf sample preservation in silica gel container.

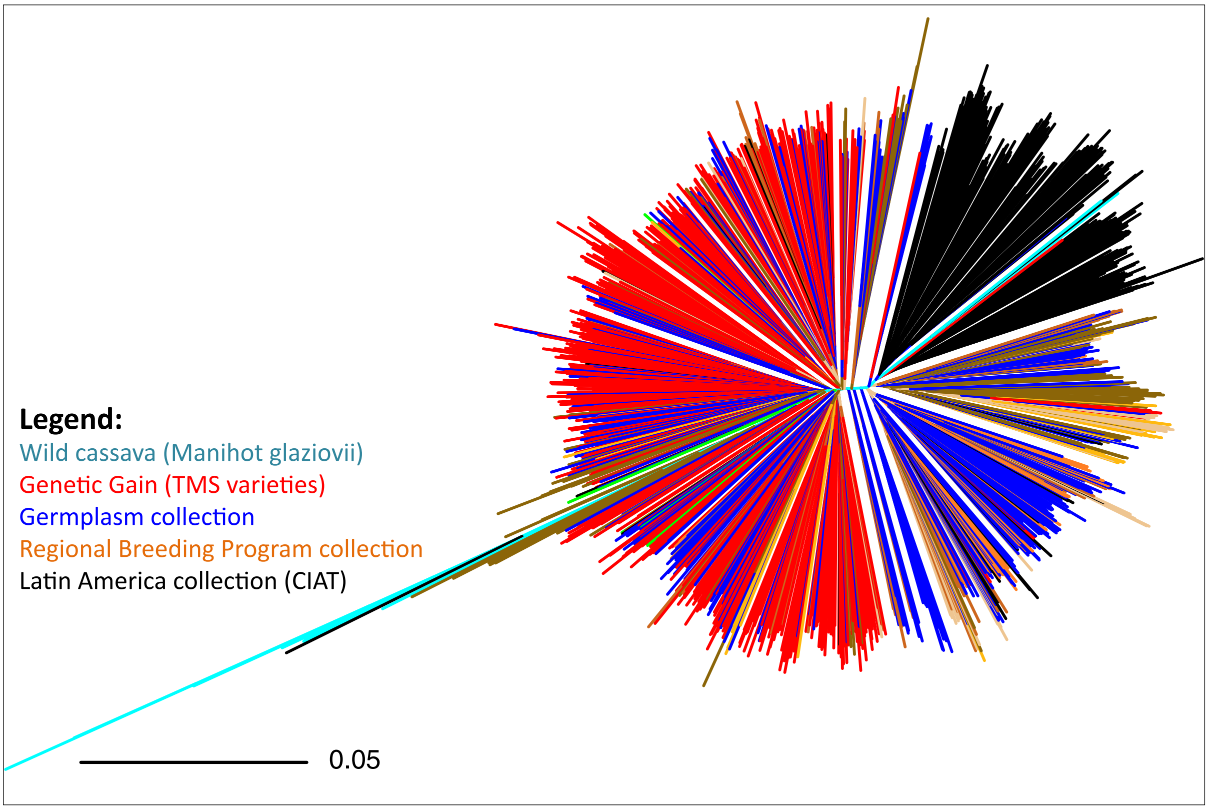


Figure 3. A neighbor-joining tree consisting of a reference library of diverse genotypes of multiple sources.

Bioinformatics

An initial raw sequence data of approximately 200 Gbp were generated after sequencing 7565 genotypes. The raw read sequences of the samples in the current study were processed in combination with the accessions in the reference library through cassava TASSEL-GBS discovery pipeline (Fig. 4) developed using TASSEL 5.0 and initially generated with about 2500 cassava clones under the NextGen Cassava project (www.nextgencassava.org). A SNP calling was performed based on TASSEL-GBS production pipeline (Fig. 5) by aligning the tags to the most recent cassava reference genome, version 6.0.

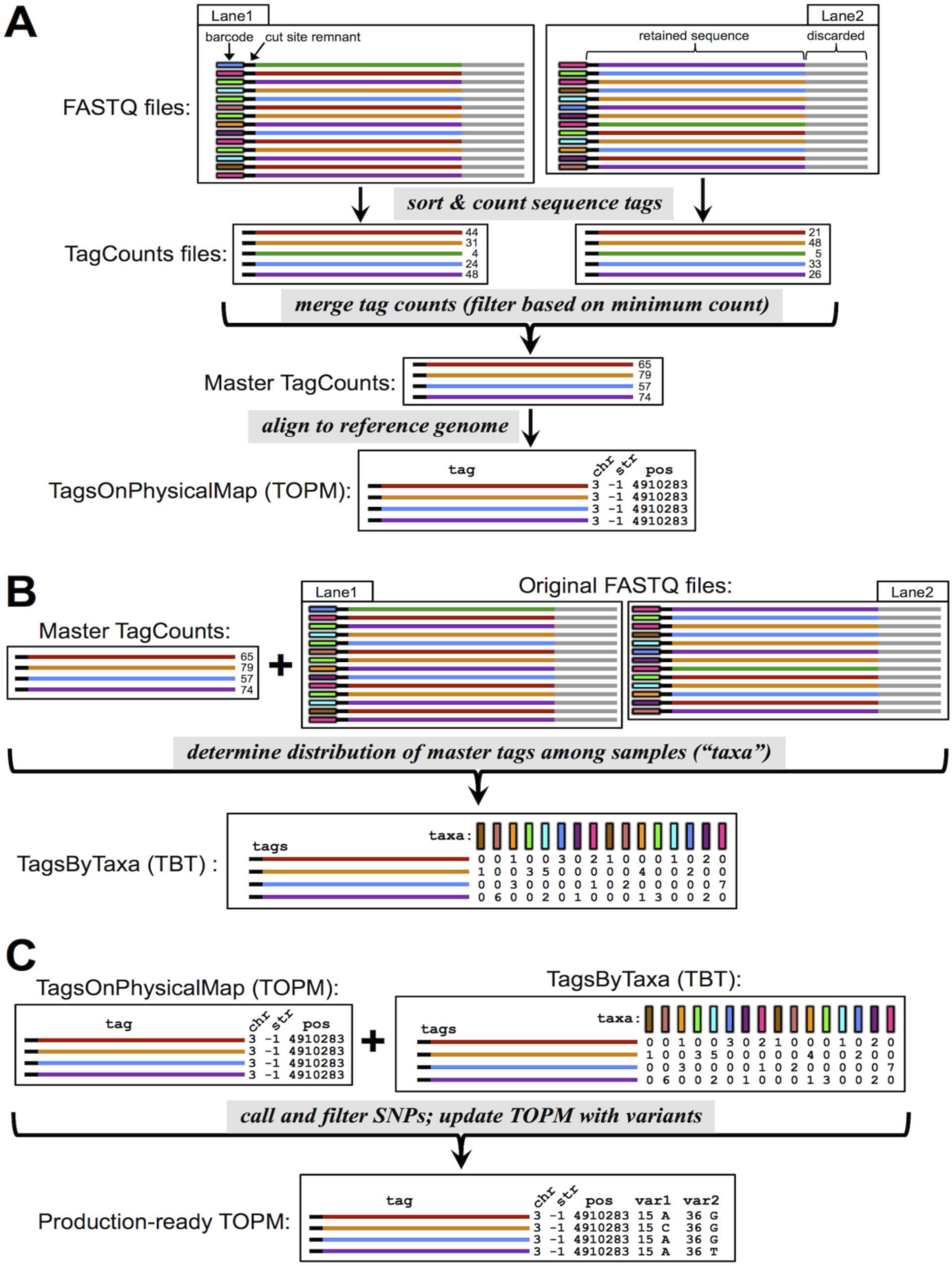


Figure 4. Schematic representation of the TASSEL-GBS Discovery Pipeline adapted from Glaubitz et al. 2014.
(A) Barcoded sequence reads are processed and collapsed into a set of unique sequence tags, with one TagCounts file produced per input FASTQ file. The separate TagCounts files are then merged to form a "master" TagCounts file, which retains only those tags present at or above an experiment-wide minimum count. This master tag list is then aligned to the reference genome and a TagsOnPhysicalMap (TOPM) file is generated, containing the genomic position of each tag with a unique, best alignment. (B) The barcode information in the original FASTQ files is then used to tally the number of times each tag in the master tag list is observed in each sample ("taxon") and these counts are stored in a TagsByTaxa (TBT) file. (C) The information recorded in the TOPM and TBT is then used to discover SNPs at each "TagLocus" (set of tags with the same genomic position) and filter the SNPs based upon the proportion of taxa covered by the TagLocus, minor allele frequency, and inbreeding coefficient (FIT). For each retained SNP, the allele represented by each tag in the corresponding TagLocus is recorded in the TOPM file, along with its relative position in the locus. The end product of the Discovery Pipeline is a "production-ready" TOPM that can then be used by the Production Pipeline to call SNPs.

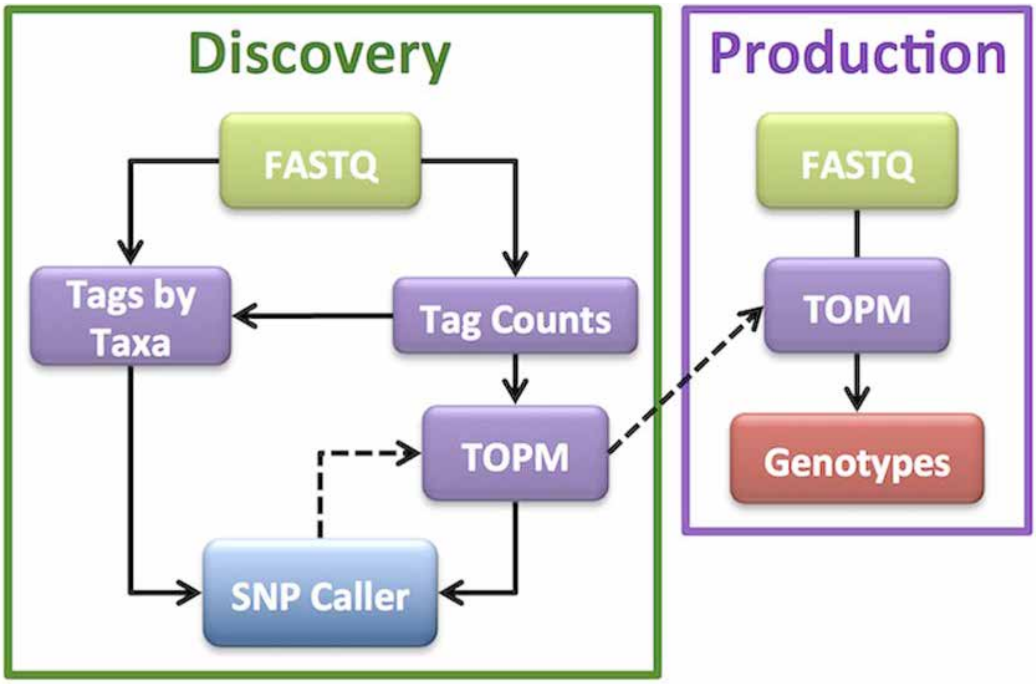


Figure 5: The TASSEL-GBS production pipeline and its relationship with discovery pipeline. Adapted from Glaubitz et al. 2014.

The Discovery Pipeline is run periodically on all FASTQ files generated to date in a species, and the ascertained and filtered SNPs are stored in a "production-ready" TOPM. The Production pipeline utilizes this production-ready TOPM to quickly call SNPs either for the original samples in the Discovery Build, or for subsequent, post-Discovery samples.

Resulting hapmap files (SNPs) were filtered for quality control of SNP data by removing those with > 30% missing data and minor allele frequency of < 0.01%. A total of 62,548 single nucleotide variants (SNV) were initially discovered from 11,578 accessions consisting of 3891 of the reference library samples and 7411 of the CMS samples. Filtering as a quality control measure was done by missing genotype rates at maximum per-variant == 0.6 and maximum per-sample == 0.8. A total of 276 accessions and 9649 SNV were removed owing to missing data. The final data therefore consisted of 52,899 variants and 11,302 accessions that passed filters and quality control.

Distance threshold to establish criteria to determine identical sets of genotypes

A total of 89 samples randomly selected were genotyped in duplicate to determine a distance threshold between genotypes that could help to declare a distance at which two or a set of genotypes are similar or distinct. A frequency distribution of distance (IBS) was plotted and resulted in bimodal distribution of pairwise genetic distance. The bimodal distribution shows the frequency distribution of the data where one of the curves shows an artefact that could occur owing to genotyping error. Filtering with minor allele frequency (MAF) at 0.00025, missing individuals at 0.8 and missing genotypes of 0.6 was done and resulted in a total genotyping rate of 0.642083 with total variants of 35729. A frequency histogram of distance then resulted in a distance threshold value of 0.0166. The red dashed line at 0.0166 as indicated on the dendrogram was therefore declared as a distance threshold where any pair of genotypes or set of genotypes below the line or less than 0.0166 are identical (Fig. 6).

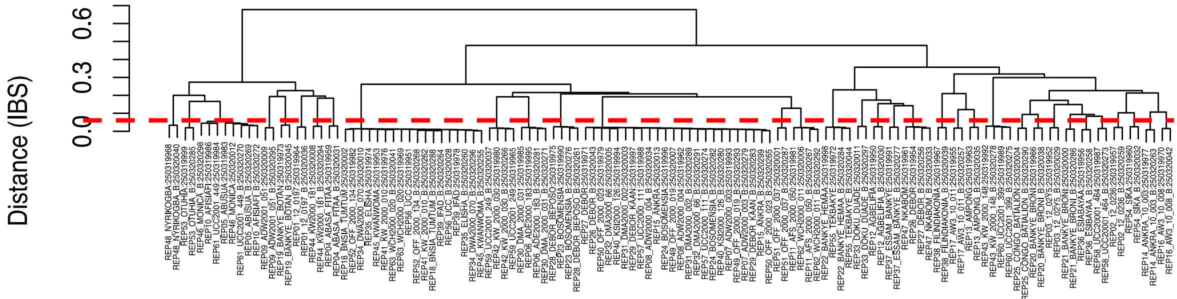


Figure 6. Established criteria for determining identical sets of clones based on 89 samples genotyped in duplicates.

Cluster analysis

Identification of cassava varieties collected from farmers' fields was first performed using distance-based hierarchical clustering, a pairwise genetic distance (identity-by-state, IBS) matrix calculated from 52,899 SNP markers in PLINK: Whole genome data analysis toolset version 1.9 (Purcell et al., 2007). A Ward's minimum variance hierarchical cluster dendrogram was then built from the IBS matrix using the Analyses of Phylogenetics and Evolution (ape) package in R software. The critical distance threshold to declare whether two genotypes are identical was empirically determined from the distribution of pairwise distances between duplicated DNAs from 89 samples. This "calibration principle" approach was taken because of the possibility of SNP genotype errors resulting from miscalling as homozygotes some heterozygous SNPs with low sequencing read depth.

Matching farmers' varieties to those in the reference library

The hierarchical clustering of 11302 accessions based on Ward's minimum variance revealed two main cluster groups representing the genetic gain collection and IITA germplasm collections (Figs 7 and 8). The genetic gain and germplasm collections have also prefix names: TMS representing breeding lines and TME representing the gene bank collection. Individual accessions in the same cluster with the distance threshold below 0.0166 were all considered as the same genotypes. Any genotype from the reference library that falls in the cluster of different individuals representing the same variety, is identified based on the reference library. Of the total 7376 farmers' varieties collected in the current study only 4822 matched genotypes in the reference library, whereas 2554 did not match any of the varieties there. On the other hand, 1663 of the 3891 genotypes in the reference library did not match any of the varieties in CMS. Those varieties not matching the genotypes in the reference library were further observed for any match to the genetic gain cluster groups so that their improvement status could be considered. Altogether a total of 114 different varieties were identified (Table 25). Among these, 46 varieties matched the genetic gain cluster group whereas 68 matched landrace groups, 18 matched officially released varieties, and 15 matched varieties that are improved and released. Among the officially released varieties only 14 were from the genetic gain group whereas the remaining four cultivars represented landrace collections evaluated on experimental plots and officially released.

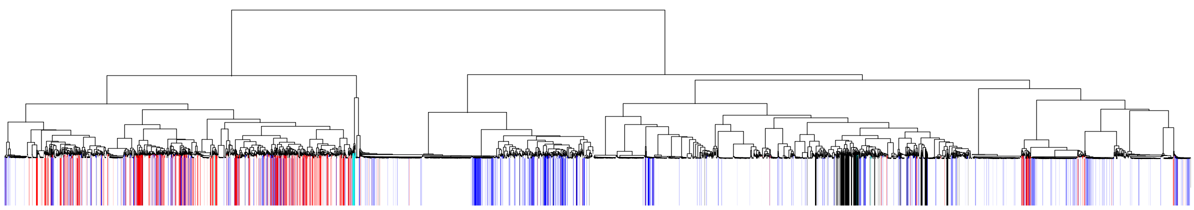


Figure 7. The TMS versus TME cluster groups. The red bars represent the genetic gain whereas the blue bars represent IITA germplasm collection.

Figure 8. Hierarchical cluster of 11,302 accessions representing both the CMS collection and the reference library.

Table 25. List of genotypes in the reference library matching farmers' field collected varieties with its respective genetic group, release information and improvement status.

Matched cultivar	Classification (GG or LR)	Matched released varieties	Status [improved/improved + released/released]
TMEB1116/TMEB99/TMEB1135/ TMEB1105/TMEB1099/TMEB1212	GG	NO	Improved
TMSI920326	LR	TMS 92/0326	Improved + Released
TMEB597	LR	NO	NO
AKPU/TMEB1142	LR	NO	NO
TMEB2/ODONGBO	LR	MS-3 (Odongbo)	Released
TMS30572	GG	TMS-30572	Improved + Released
IITA_GRU_LR_core_258	GG	NO	LR + Match GG
TMEB966/TMEB965	LR	NO	NO
TMSI980581	GG	TMS 98/0581	Improved + Released
TMEB419	LR	TME-419	Released
TMEB1044/TMEB1042	LR	NO	NO
TMEB1641/TMEB265/TMEB1643/ TMEB1992/TMEB1774	LR	NO	NO
TMSI9102324	GG	NO	Improved
TMSI010760	LR	NO	Improved
TMEB104/TMEB102/TMEB37	LR	NO	NO
TMEB971/TMEB1736	LR	NO	NO
I9102325	GG	NO	Improved
TMEB3, 4,7,12,14	LR	NO	NO
ANTIOTA/TMEB1	LR	MS-6 (Antiota)	Released
TMEB1369/TMEB634/TMEB813	LR	NO	NO
CW5251/NR8082	GG	NR-8082	Improved + Released
BEN 86019/Caricass/Toma	LR	NO	NO
TMEB1040/TMEB1050	LR	NO	NO
TMEB117/ANKRA	LR	NO	NO
TMSI980505	GG	TMS 98/0505	Improved + Released
TMSI071313	GG	NO	Improved
MBOMA11	LR	NO	NO
TMEB491/TMEB508	LR	NO	NO
TMEB1858/TMEB119/TMEB1024	LR	NO	NO
TMSI011368	GG	IITA TMS 1011368	Improved + Released
TMSI4(2)1425	GG	TMS-4(2)1425	Improved + Released
GRU_LR_core_157	LR	NO	NO
TMEB262/TMEB1647/TMEB410/ TMEB2048	LR	NO	NO
TMEB1000	LR	NO	NO
TMSI940270	LR	NO	Improved
NR7734	GG	NO	Improved
TMEB120	LR	NO	NO
Malawi_034	GG	NO	LR + Match GG

Matched cultivar	Classification (GG or LR)	Matched released varieties	Status [improved/improved + released/released]
TMEB33	LR	NO	NO
TMEB621	LR	NO	NO
TMEB47	LR	NO	NO
TMSI30040	LR	NO	Improved
TMEB8	LR	NO	NO
TMEB34	LR	NO	NO
TMEB131 (C-2 Cape Verde) /TMEB2050/ TMEB994	LR	NO	NO
TMSI011412	GG	IITA TMS 1011412	Improved+ Released
AR1-81/CR12-45/AR37-108 (TME961-WarriNorthDelta)	LR	NO	Improved
GRU_LR_core_198	LR	NO	NO
TMEB470/TMEB625/TMEB2089/ TMEB2056/TMEB933	LR	NO	NO
Cameroon_010	LR	NO	NO
Cameroon_030	LR	NO	NO
TMEB10	LR	NO	NO
MOCUBA	LR	NO	NO
TMS50395	GG	TMS-50395	Improved + Released
TMEB1818	LR	NO	NO
TMEB463/235	LR	NO	NO
TMEB1200/TMEB983	GG	NO	LR + Match GG
TMEB1485/TMEB1278	LR	NO	NO
TMSI920057	GG	TMS 92/0057	Improved + Released
W940727	GG	NO	Improved
TMEB2122	LR	NO	NO
TMEB945	LR	NO	NO
TMSI000354	GG	NO	Improved
TMEB9	GG	NO	LR + Match GG
Ug130005	GG	NO	LR + Match GG
Ug120078	LR	NO	NO
TMSI011752	LR	NO	Improved
TMSI30555/NR8212	GG	TMS-30555	Improved + Released
TMSI960325	GG	NO	Improved
TMSZ010087	LR	NO	Improved
TMEB981	LR	NO	NO
TMSI982101	GG	NO	Improved
Ug120140	GG	NO	LR + Match GG
Ug120142	GG	NO	LR + Match GG
TMSI972205	GG	TMS 97/2205	Improved + Released
B9200068	LR	NO	Improved
TMSI011371	GG	IITA TMS 1011371	Improved + Released
SLICASS3	LR	NO	NO
TMEB469/TMEB468/TMEB2023/ TMEB2021/TMEB2068	LR	NO	NO
TMSI30211/I40764	GG	NO	Improved

Matched cultivar	Classification (GG or LR)	Matched released varieties	Status [improved/improved + released/released]
TMEB148	GG	NO	LR + Match GG
TMSI980002	GG	TMS 98/0002	Improved + Released
TMEB286	LR	NO	NO
TMEB1386/TMEB1484/TMEB1509	LR	NO	NO
TMEB1321	GG	NO	LR + Match GG
CR14B218	GG	NO	Improved
TMSI020452	GG	NO	Improved
TMEB499	LR	NO	NO
TMEB23/TMEB24	LR	NO	NO
TMEB279/TMEB624	LR	NO	NO
TMSI60506	GG	NO	Improved
TMEB121	LR	NO	NO
TMEB1019/TMEB1002/TMEB1046/ TMEB1021/TMEB1006	LR	NO	NO
TMEB30/TMEB31/TMEB21	LR	NO	NO
AR9-19	LR	NO	Improved
TMSI960304/TMSI960804	LR	NO	Improved
TMEB1001	LR	NO	NO
COB1139	GG	NO	Improved
TMEB2121	GG	NO	LR + Match GG
TMSI9102322	LR	NO	Improved
TMSI974779	GG	NO	Improved
TMEB497/TMEB1025	LR	NO	NO
I083389/I940561	GG	NO	Improved
B9200061	LR	NO	Improved
NR050080	GG	NO	Improved
CW45113	LR	NO	Improved
TMSI011646	GG	NO	Improved
NR090146/NR110122/NR87184	GG	NR 87184	Improved + Released
TMSI011811/I010536	GG	NO	Improved
NZIVA	GG	NO	LR + Match GG
TMEB762/TMEB757	LR	NO	NO
TMEB960	LR	NO	NO
TMEB952	GG	NO	LR + Match GG
Ug120145	GG	NO	LR + Match GG

GG = genetic gain; LR = landrace; LR + Match GG = varieties matching landrace in the genetic gain group.

Extent of improved and/or released varieties on farmers' field across all samples investigated based on DNA fingerprinting

In addition to the level of genetic matching of the farmers' varieties to the genotypes in the reference library it is also important to have variety release information. This will help to have clear figures on further analysis on the adoption of improved varieties based on DNA fingerprinting. In the current study more than 35% of the total number of samples collected on the farmers' fields represented "improved varieties" which are categorized in three different groups: improved and officially released,

Table 26. Percentage of improved and/or released varieties on the farmer's field based on DNA fingerprinting.

Varieties	GG	LR	Total	%	Cumulative %
Improved AND released	919	3	922	12.50	12.50
Improved NOT released	267	491	758	10.28	22.78
Single accessions in GG cluster	110		110	1.49	24.27
Not in Library, in GG cluster	598		598	8.11	32.38
Matching LR, in GG cluster	55		55	0.75	33.12
Released but TMEB1, TMEB2..		926	926	12.55	45.68
Local varieties		4007	4007	54.32	54.32
Total	1949	5427	7376	100.00	

GG=genetic gain; LR=landrace;

improved but not officially released, and released but matching those accessions from the landrace (TME cluster group). Results (Table 26) also show that about 33% of the varieties matched the genetic gain cluster group. However, only 12.50% matched improved and officially released varieties. In addition, 12.56% of the cultivars from farmers' fields matched those landrace collections that went through field evaluation and official release. A significant percentage (10.28) of the farmers' varieties also represent improved varieties in the GG group that were not officially released but made their way to farmers' fields.

Singular varieties not matching any of the genotypes but found in the GG group represent 1.49% of the collection. Similarly, about 8.11% represent varieties matching the GG cluster group but none were in the reference library. The larger proportion (54.32%) represents local varieties.

Farmers' variety matching released varieties

IITA in collaboration with national agricultural research systems (NARS) has officially released 46 cultivars. However, in the current study conducted in the main cassava growing regions of Nigeria a large number (28) of the 46 cultivars were not observed on farmers' fields, suggesting that these were either disadopted or no initial dissemination effort had been made (Table 27). Among the 28 varieties not found on farmers' fields 11 were not in the reference library, whereas none of the varieties matched the remaining 17, which are in the library. Of the officially released varieties only 18 were encountered on farmers' fields and five were the dominant varieties growing in the study regions with a frequency greater than 100 (Fig, 9). These are TMS30572, MS-3 (Odongbo), MS-6 (Antiota), TMS 50395, and TME 419 in descending order of their frequencies on farmers' fields. Seven of them are very rare, with a frequency less than ten.

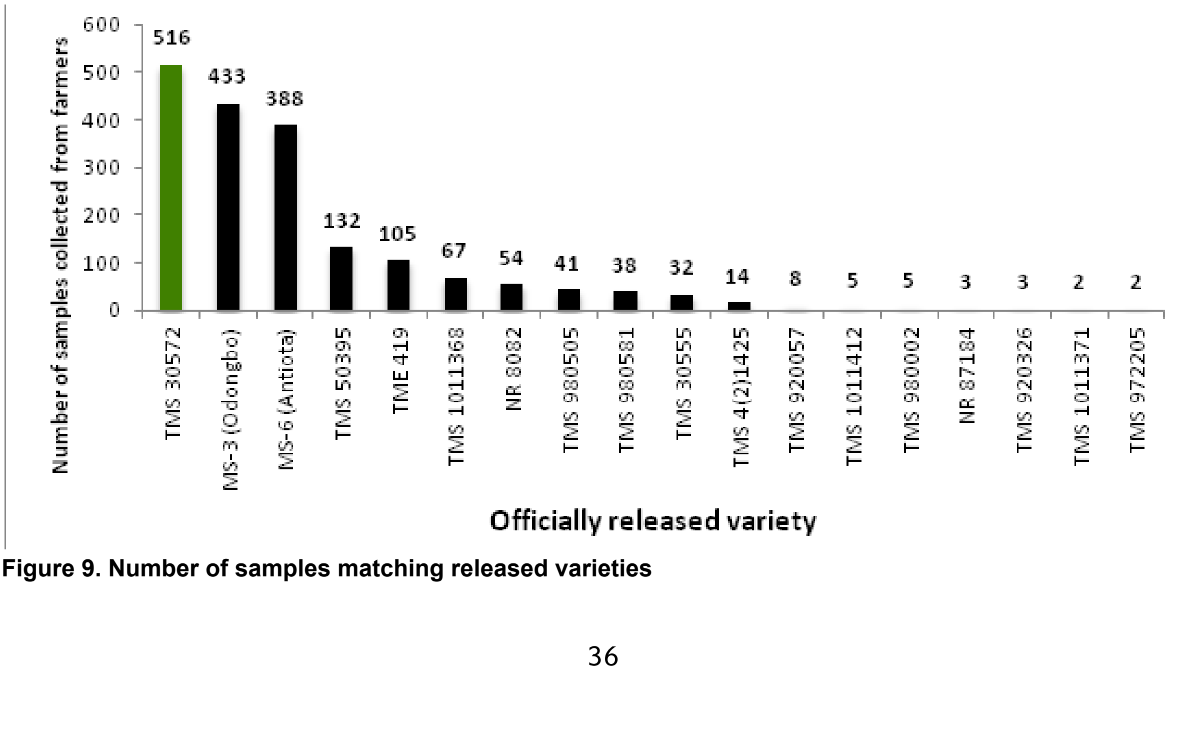


Figure 9. Number of samples matching released varieties

Table 27. List of released cassava varieties in Nigeria and matching with the current collection.

TMS-30572	Found in CMS
TMS-4(2)-1425	Found in CMS
MS-6 (Antiota)	Found in CMS
MS-3 (Odongbo)	Found in CMS
TMS-30555	Found in CMS
NR-8082	Found in CMS
TMS-50395	Found in CMS
TME-419	Found in CMS
TMS 97/2205	Found in CMS
TMS 98/0505	Found in CMS
TMS 98/0581	Found in CMS
NR 87184	Found in CMS
TMS 92/0057	Found in CMS
TMS 92/0326	Found in CMS
TMS 98/0002	Found in CMS
TMS 10/11368	Found in CMS
TMS 10/11412	Found in CMS
TMS 10/11371	Found in CMS
TMS 82/00058	Only in Library
NR 8212	Only in Library
TMS 91934	Only in Library
TMS 98/0510	Only in Library
TMS 96/1632	Only in Library
NR 93/0199	Only in Library
TMS 96/1089A	Only in Library
NR 01/0004	Only in Library
CR 41-10	Only in Library
TMS 01/0040	Only in Library
TMS 00/0203	Only in Library
NR 03/0211	Only in Library
CR 36-5	Only in Library
TMS 98/2132	Only in Library
TMS 01/1206	Only in Library
TMS 07/0593	Only in Library
TMS 07/0539	Only in Library
TMS-90257	Not in Library
TMS-84537	Not in Library
TMS-82/00661	Not in Library
TMS-81/00110	Not in Library
NR-8208	Not in Library
NR-8083	Not in Library
NR-83107	Not in Library
NR-41044	Not in Library
TMS-30001	Not in Library
NR 03/0155	Not in Library
NR 07/0220	Not in Library

Frequency of varieties in the collection

The majority of the varieties were less frequent (<20 times) on farmers' fields across the cassava-growing regions. However, 16 cultivars were the most common varieties with a frequency of at least 100 times (Figs 10 and 11) ranging from the least 119 for TMEB499 to 516 for TMS30572, whereas 206 accessions occur only once in the collection. Among the 18 officially released varieties observed on farmers' fields only five were dominant varieties growing in the study regions with a frequency greater than 100 (Fig.10). The remaining varieties were less frequent with seven of them occurring less than 10 and the other six ranging between 32 and 67.

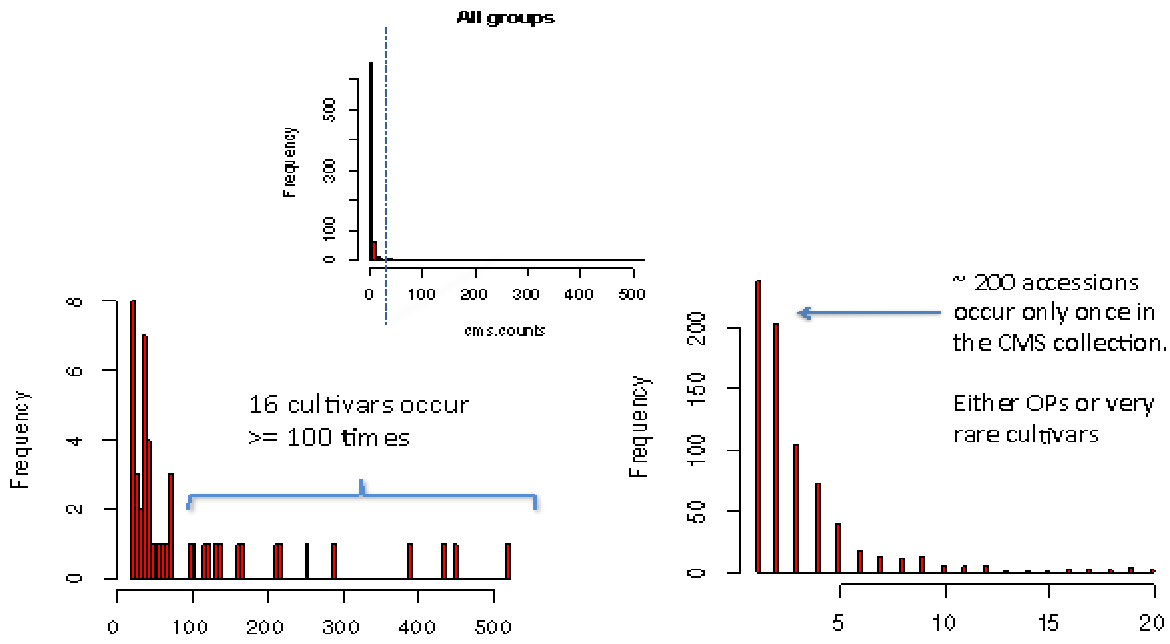


Figure 10. Major and minor cultivars growing in the four study regions

Number of varieties

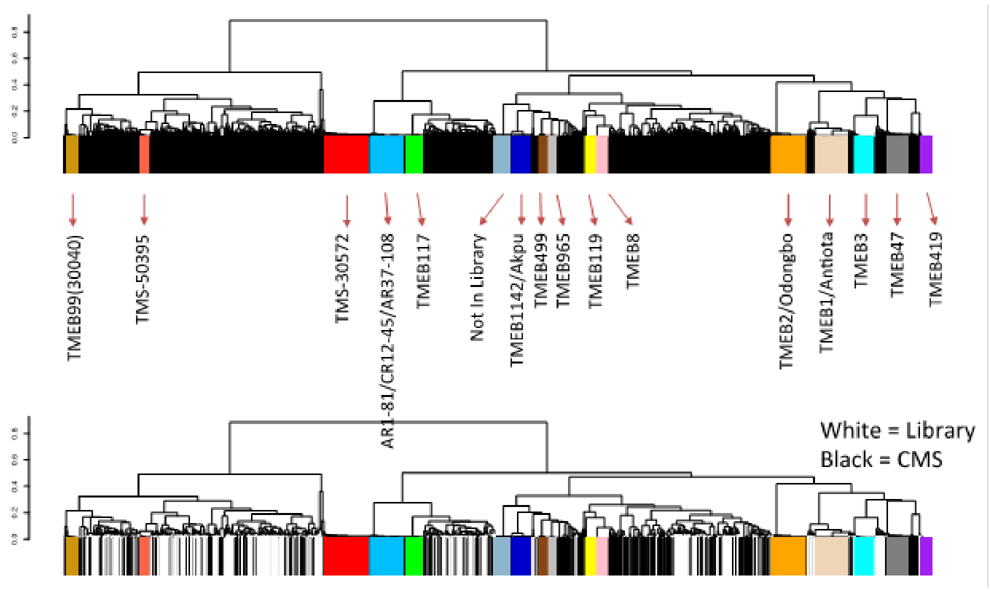


Figure 11. Major varieties with frequency of greater or equal to 100

Geographical distribution of the identified varieties

Different patterns of distribution of varieties were observed across the study regions. Among the total 114 different varieties identified, 21 were sampled from all four regions, 14 in three regions, 19 in two regions, and 60 were growing only in one of the four regions (Fig. 12). The highest number of varieties (65) was found in the Southwest followed by North Central (61), South-South (54), and Southeast (44) (Fig 13). There is also variation in terms of the magnitude in which varieties grow in the different regions. For example, among the officially released varieties TMS30572, TMS50395, and MS-6 (Antiota) are dominant in the Southwest whereas AR1-81/CR12-45/AR37-108 are dominant in the South-South and MS-3 (Odongbo) in the North Central region (Fig. 14).

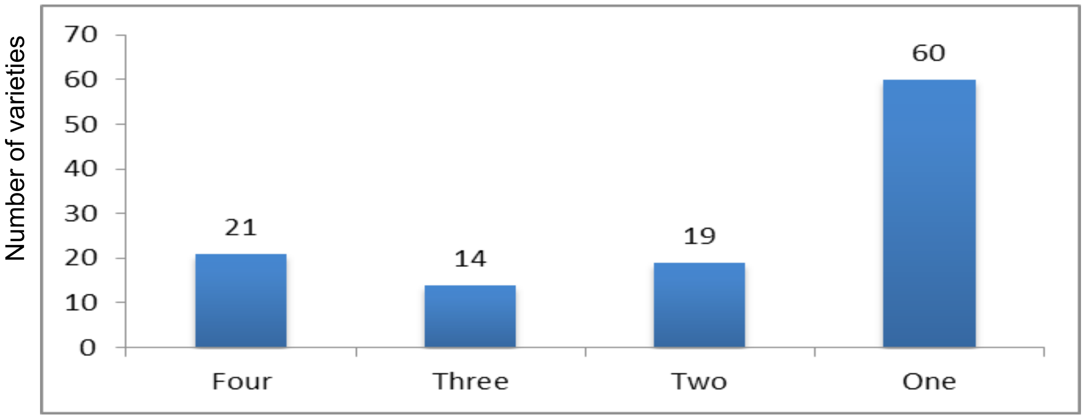


Figure 12. Number of varieties growing in all the regions, three of the regions, two of the regions and only in one of the four regions.

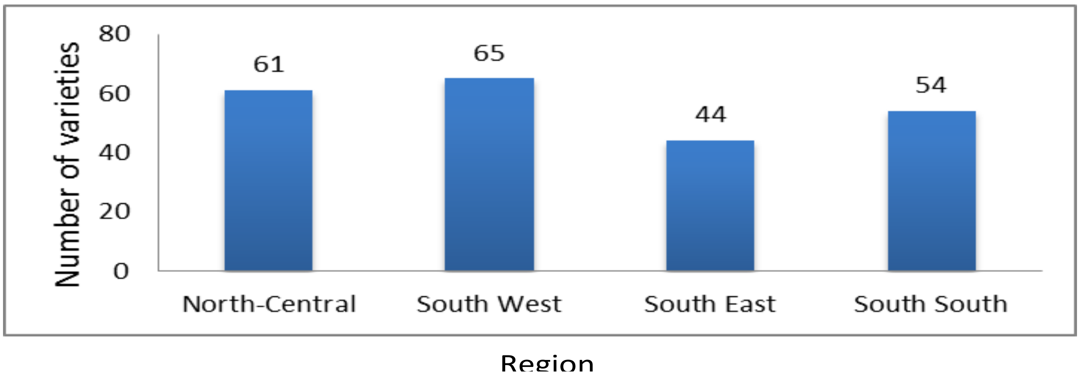


Figure 13. Number of different varieties growing across the four study regions.

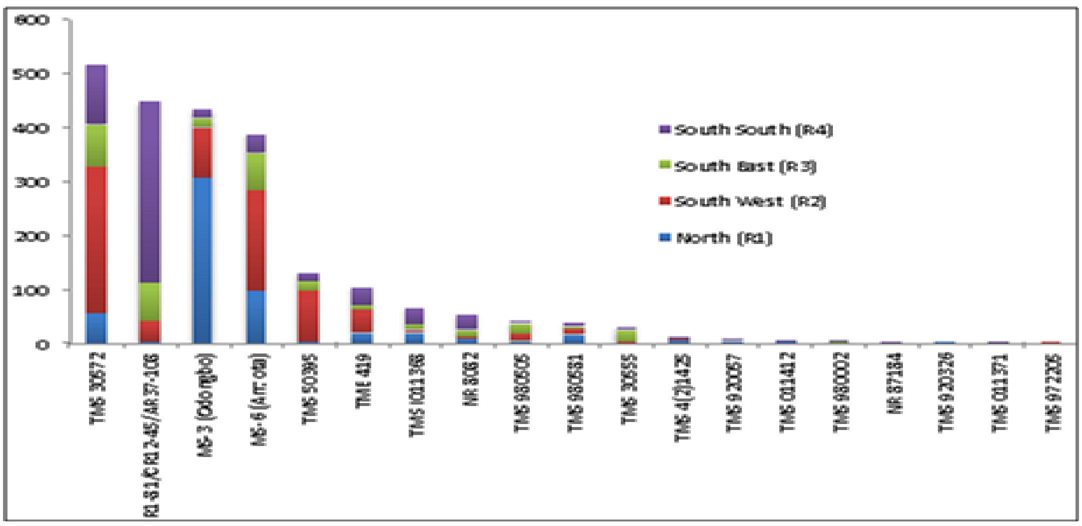


Figure 14. Frequency and geographic distribution of cultivars across the study regions

Comparison of adoption rates based on matched self-reported and DNA data

Take-home message

- We found that, on average, adoption rates are similar when using DNA fingerprinting and self-reported data.
- Even though the average adoption rate seems similar between self-reported and DNA fingerprinted data, misclassification rate is large.
- For examining the determinants of adoption, measurement of "improved varieties" is relevant as the determinants of adoption are different while using self-reported and DNA fingerprinted data.
- Three sources of heterogeneity largely explain the probability of correctly classifying varieties into improved and local varieties. These include level of education, access to information (such as mobile phone ownership and to extension), membership in formal and informal organizations and location.

In this section, we matched farmers own self-reported and DNA based varietal identification to identify the extent of adoption of improved cassava varieties.

Extent of adoption of improved varieties based on matched CMS-DNA data

CMS collected data on the adoption rate of improved varieties at varietal level, the lowest aggregation point. Using these data, we calculated the adoption rate at the household level as a percentage of households that grow at least one improved variety. In addition, we also calculated adoption rates at each aggregation level based on farmers' self- reported data as well as DNA-based varietal identification. The result shows that adoption at the household level is close to 60% for Nigeria as a whole (i.e., about 60% grow at least one improved variety in one or more plots) when using farmers' self-reported data. Figure 15 reports the adoption rate for the full sample as well for the regions considered in the CMS study. Southwest, one of the leading cassava processing regions, has the highest self-reported adoption rate (79%[6]). Figure 15 also reports the adoption rate using DNA-based varietal identification. However, classifying varieties into improved and landraces is not

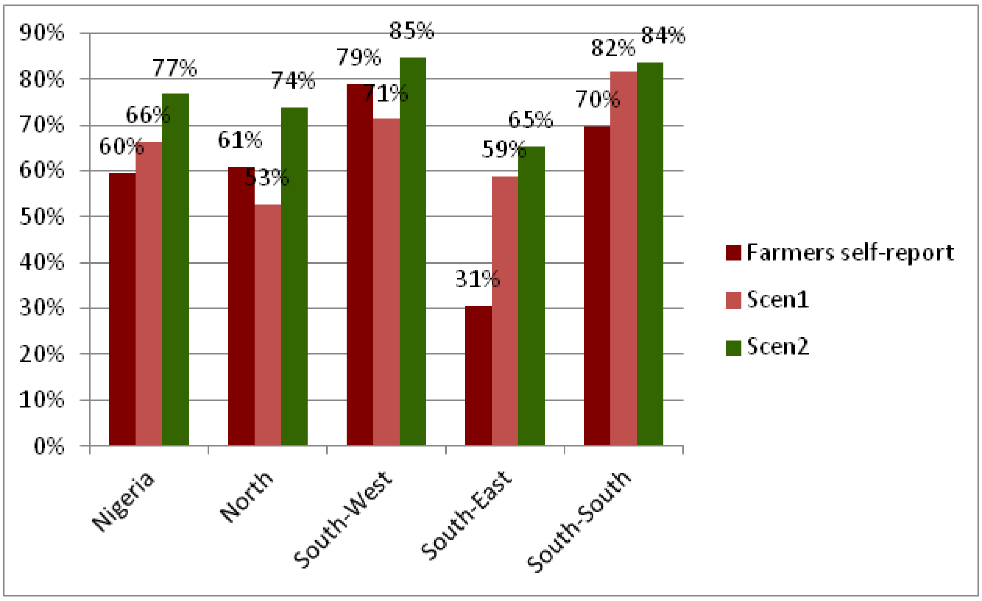


Figure 15. Adoption rate based on matched CMS-DNA data

[6]IITA is located in this region.

Table 28: Categories used for defining improved varieties based on DNA fingerprinting results.

Serial no.	Variety Groups	Definition
1	Improved and Released	This category includes varieties that match improved and released varieties in the reference library.
2	Improved not Released	Again, this group of varieties matches improved varieties (i.e. clones in genetic gain) but are not in the officially released list. Varieties do go out before release because of on-farm trials, pre-release demonstration and other non-official channels.
3	Single accessions in GG cluster	These are accessions that genetically do not match any specific clone in the collection but cluster with improved varieties in the reference library. This can happen when cross-pollinated (or even self-pollinated) seeds germinate in farmers' fields and the farmers eventually propagate these as a variety.
4	Not in Library, in GG cluster	No library can completely encompass every possible genotype found in household farms. As this is a clonal crop, the breeding program cannot maintain every genotype that it produced or even disseminated because of natural attrition over time. So where we found clones that did not match any specific genotype in the library, the next best thing was to use the cluster analysis method and find where they belong. In this case, these accessions clustered with the improved variety collection we have in the library.
5	Matching LR, in GG cluster	There is an explanation for this group. The "landraces" (LR) are accessions whose name start with prefix "TMEB" and were gathered during germplasm collection expeditions in Nigeria and other African countries over the years. Many of these are obviously landraces but some, in reality, are improved varieties which had lost their original names, were collected again, and brought back to IITA. Because of their unknown identities, these accessions are usually placed in the landrace collection, even though some are improved.
6	Released but TMEB1, TMEB2	Among the officially released varieties are some genotypes that were not developed through formal breeding processes (i.e., not from breeder crosses). After several years of purification and testing, these landraces were found to have superior characteristics and were recommended for release. In some instances these genotypes were transferred from one country/region to another which would not have occurred without the intervention of formal breeding programs like IITA. A good case is Gbazekuote (TMEB419) which was brought from Togo and is now grown in Nigeria and other countries. It is an officially released variety.
7	Local varieties	This are landrace varieties

straightforward even after DNA fingerprinting due to issues of measurement and library matching. We therefore develop two scenarios based on commonly used criteria for the identification of varieties into improved and non-improved. Table 28 presents the varietal groups as well as definitions for each group as obtained from DNA fingerprinted results.

Based on this information, we constructed two scenarios for defining adoption rate:

Scenario 1: Adoption based on improvement status:
In this scenario, the distinction between improved variety and landrace will be defined, based on improvement status. It includes nos 1, 2, 3, 4, 5 and TMB419[7] from category 6 from Table 28.

Scenario 2: Adoption based on release information and genetic gain
This scenario defines improved varieties based on improvement status as in Scenario 1 but includes landraces that were released after some purification. It includes case 1, 2, 3, 4, 5, and 6 from Table 28. Adoption rate for each scenario is presented in Figure 15. The result suggests that, on average, adoption rates are similar when using DNA fingerprinted and self-reported data except for Scenario 2 in which landrace varieties that were released after some purification were considered as improved varieties.

[7]TMB419 is included under improved varieties as it is imported from Togo.

For instance, according to farmers' self-reported data, the adoption rate at the household level stands at 60%. When DNA-based varietal identification is used, the adoption rate ranges from 66 to 77%, depending on the scenario used for categorizing varieties into improved and local. In what follows, we also report the adoption rate based on intensity of adoption at the plot level (Fig. 16). Intensity of adoption is calculated by considering the area under improved varieties out of the total cassava area. The result shows that despite high rates of adoption the intensity of adoption is very low. The current adoption rate, based on intensity of adoption, stands at 38% while using farmers' self-reported data. Regional distribution of adoption rates further reveals that the adoption rate is highest in the Southwest region. The lowest adoption rate is reported in the Southeast. These results are not surprising as IITA is located and has been operating in the south-western part of the country for the last 50 years. The South-South region of the country has the second highest adoption rate. This might be due to the presence of national research centers in the region[8].

However, the above average values mask a lot of heterogeneity because of errors of classification. In statistics, this misclassification is commonly referred to as false negatives (considering something false when it is true) and false positives (considering something true when it is false). Based on our matched self-perception and DNA fingerprinted data in Fig 16, we calculated the rate of misclassification in the adoption rate. These include (a) households who think that they grow improved varieties but actually grow local varieties as obtained from the DNA fingerprinting results (Group 1), (b) households who think that they grow local varieties but actually grow improved varieties according to DNA fingerprinting results (Group 2), (c) households that grow improved varieties and confirmed by DNA fingerprinting (consistent sub-sample-group 3), and (d) households that grow local varieties and confirmed by DNA fingerprinting (consistent sub-sample-group 4). Table 29 presents misclassification rates based on household level adoption data.

Table 29 suggests that, even though the average adoption rate seems similar (between farmers'-self-reported and DNA fingerprinted data), the rate of misclassification is large. As shown (Table 29), only 66.76% in Scenario 1 and 66.75% in Scenario 2 were correctly classified and the remainder is misclassification. For instance, in Scenario 1, about 19.98% of the households think they grow local varieties but actually grow improved varieties. Similarly, 13.26% of the households think they grow improved varieties but actually grow local varieties. These misclassification rates are also quite prevalent in Scenario 2. This result implies that because of large misclassification rates, wrong policy interventions may end up being promoted for the wrong group of households.

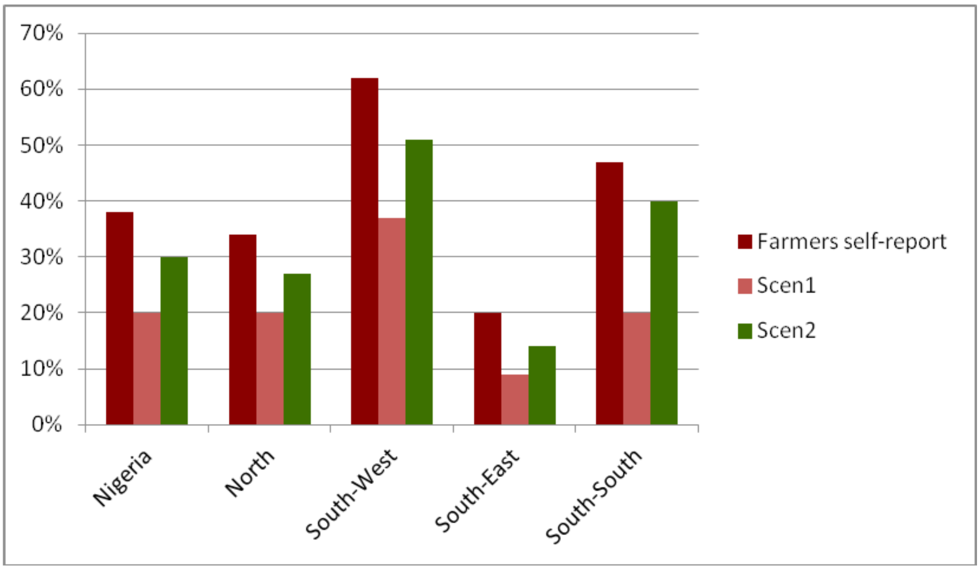


Figure 16. Intensity of adoption based on matched CMS-DNA data.

[8] Intensity rates for each region is presented in figure 21 in the appendix (intensity of adoption ranges from 62% in Southwest to about 20% in the Southeast

Table 29. Misclassification rate.

Classification	Scenario 1	Scenario 2
Adopter according to DNA and non-adopter according to self-perception	19.98%	25.22%
Adopter according to self-perception and non-adopter according to DNA	13.26%	8.03%
Adopter in both	46.34%	51.57%
Non-adopter both	20.42%	15.18%
Correctly classified (row 3+row 4)	66.76%	66.75%
Total	100%	100%

Table 30. Regional misclassification rate based on household level adoption rate.

Classification	Region	Scenario 1	Scenario 2
Adopters according to DNA and non-adopters according to self-perception	North	9.2%	21.3%
	South-west	10.7%	15.4%
	Southeast	37.1%	41.1%
	South-South	20.5%	21.4%
Adopters according to self-perception and non-adopters according to DNA	North	17.9%	8.6%
	South-west	18.4%	9.6%
	Southeast	8.9%	6.4%
	South-South	8.8%	7.6%
Adopters in both	North	43.6%	52.8%
	South-west	60.6%	69.4%
	Southeast	21.8%	24.2%
	South-South	61%	62.2%
Non-adopters in both	North	29.4%	17.3%
	South-west	10.3%	5.6%
	Southeast	32.3%	28.2%
	South-South	9.7%	8.8%

Results reported in Table 30 further confirm that misclassification rates at the regional level are also quite high. In Scenario 1, correct classification rates ranged from 64.1% in the Southeast to 73% in the North. While using scenario 2, the correct classification ranges from 52.4% in the Southeast to 71% in the South-South.

Determinants of adoption using matched CMS-DNA data

In Section 2.9, we looked into the determinants of adoption of improved cassava varieties using farmers own self-reported adoption data. In this section, we compare the results by estimating the same model albeit with DNA fingerprinting data. Undertaking this kind of analysis is crucial due to a high rate of misclassification. Table 31 presents the determinants of adoption using farmers own self-reported and DNA fingerprinting data for each scenario. The results show that some socioeconomic characteristics have robust effects across all measures of adoption. For example, household size has positive and statistically significant effect on the probability of adopting improved cassava varieties irrespective of the way improved varieties are measured. For education, we found significant effect when using DNA based varietal identification and insignificant effect when using farmer's self-reported data. Theory predicts that more educated farmers are more likely to adopt improved cassava varieties. In this regard, measuring adoption based on DNA fingerprinting data provides consistent estimates. However, for most of the socioeconomic variables, effects differ not only in magnitude but also in direction. For instance, age has a negative albeit non-linear effect on the probability of adoption when adoption is measured using Scenario 2 of the DNA fingerprinting data. When using self-reported data as well as Scenario 2, we did not find any significant effect for age. We also found the same effect for other socioeconomic variables such as sex of the household head. For example, the sex of the household head has positive and significant effect on adoption when using farmer's self-reported data. However, the sex of the household head has insignificant effect when using DNA fingerprinting data.

Even more apparent is not only the difference between self-reported and DNA fingerprinted data but also within the scenarios considered using DNA fingerprinted data suggesting that measurement is crucial. For instance, access to credit and extension have a positive and statistically significant effect on adoption when using self-reported adoption data and when adoption is defined based on Scenario 1 (where, we excluded landraces that have been released after some purification). However, in Scenario 2, access to credit has no significant effect. This result emphasizes the importance of measurement of adoption as important supply-side policy interventions (such as credit and extension access) have different effects in terms of magnitude and direction of the effects. Looking into the effect of preference heterogeneity, we again found that measurement matters. We found consistent robust effects only for quality of garri and starch content. In all models, the effect of garri quality is positive and significant on the probability of adopting improved varieties while the effect of starch content is negative and significant across all models. However, we found that the effect of other traits is contingent on the definition of improved variety. Even worse are cases where we have opposite effects for a given trait depending on the way ***improved varieties*** are defined. For example, the early maturity trait has a positive and significant effect when self-reported adoption data is used and an insignificant effect DNA fingerprinted results are used.

In terms of the determinants of intensification of improved cassava varieties, household size had a positive and significant effect on the intensification decision when DNA based varietal identification is used. This result is expected as large family size provides the necessary labor needs for intensification. However, its effect was insignificant when intensity is measured based on farmers' self-reported data. Similarly, the effect of extension access was positive and significant across all specifications. This is an interesting result given the importance of such supply side policy interventions for enhancing intensification and welfare. As expected, relaxing information asymmetry affects adoption decision only as none of the variables (ownership of mobile phones, radio and TV) are significant determinants of cassava intensification. Results further suggest that farmers are less likely to intensify growing cassava in soils of medium fertility compared with fertile soils while intensification is likely to be higher in plots managed by women.

Table 31. Determinants of adoption.

	Based on farmers self-report		Based on scenario 1		Based on scenario2	
Variable	Coefficient	Z value	Coefficient	Z value	Coefficient	Z value
Adoption (1= yes and 0, otherwise)						
Household size	0.020**	2.1	0.018*	1.78	0.021**	2.18
Education	0.001	0.25	0.011**	2.43	0.012***	2.73
Age	-0.005	-0.5	-0.020**	-2.03	-0.011	-1.12
age2	0.000	0.77	0.0001**	2.36	0.000	1.42
Sex	0.164*	1.87	-0.004	-0.04	0.065	0.73
TLU	0.000	-0.17	-0.002	-1.03	0.001	0.66
Access to credit	0.146***	3.16	0.159***	3.22	0.063	1.37
Access to extension	0.176***	3.55	0.178***	3.43	0.190***	3.89
Distance from market	-0.003	-0.72	0.002	0.45	-0.003	-0.63
Availability of cassava stems	0.173***	3.45	0.101*	1.92	0.075	1.51
Radio ownership	0.020	0.35	0.023	0.37	0.052	0.89
Mobile phone ownership	0.726***	5.55	0.605***	3.91	0.540	4
Membership in Credit and Saving groups	0.146***	3.22	0.013	0.28	0.072	1.59
Membership in cooperatives	0.120**	2.44	0.230***	4.67	0.143***	2.98
Poor soil	0.103	0.72	0.169	1.12	0.157	1.1
Medium soil	-0.099**	-1.98	-0.075	-1.38	-0.070	-1.39
Men managed plots	0.143*	1.84	0.094	1.1	0.075	0.95
Men and women managed plots	0.225***	2.74	0.045	0.5	0.139*	1.67

	Based on farmers self-report		Based on scenario 1		Based on scenario2	
Variable	Coefficient	Z value	Coefficient	Z value	Coefficient	Z value
Poundability	-0.018	-0.78	-0.029	-1.24	0.029	1.32
Quality of garri	0.199***	5.67	0.163***	4.03	0.148***	4.08
Quality of fufu	-0.056	-1.6	0.027	0.75	-0.009	-0.27
Ease in loading	-0.023	-1.07	0.026	1.16	0.012	0.56
Starch content	-0.067***	-3.1	-0.095***	-4.28	-0.061***	-2.86
Tapioca taste	0.086***	3.61	0.022	0.9	0.047**	2.05
Market demand	-0.012	-0.35	0.058	1.58	0.043	1.25
Root size	0.066	1.46	-0.042	-0.88	0.007	0.15
White color	-0.015	-0.58	-0.007	-0.28	0.002	0.1
Stores well	0.022	0.74	-0.014	-0.47	0.034	1.17
Root yield	0.109***	2.21	0.151***	2.79	0.098**	1.98
Pest resistance	-0.016	-0.62	-0.026	-0.99	-0.014	-0.54
Stem yield	-0.067***	-2.76	-0.024	-0.95	-0.066***	-2.76
Early maturity	0.098***	2.69	-0.028	-0.74	0.042	1.15
Yellow roots	0.012	0.58	0.032	1.5	0.000	0.01
Intensity of adoption						
Household Size	-0.002	-0.8	0.009**	2.09	0.002	0.55
Education	-0.001	-0.96	0.002	0.84	0.002	1.06
Age	0.011***	3.51	0.002	0.6	0.011***	3.28
age2	0.0001***	-3.63	0.000	-1.18	-0.0001***	-3.42
Sex	0.052*	1.66	-0.030	-0.69	0.022	0.63
TLU	0.000	0.16	0.000	0.47	0.000	-0.08
Access to credit	0.00002	0.01	-0.065***	-2.98	-0.022	-1.29
Access to extension	0.032**	2.03	0.066***	2.96	0.017	0.96
Distance from market	0.003*	1.85	0.004*	1.66	0.001	0.85
Availability of cassava stems	0.053***	3.28	0.032	1.48	0.055***	3.08
Radio ownership	-0.016	-0.85	0.000	-0.01	0.001	0.04
Mobile phone ownership	-0.082	-1.56	-0.084	-1.07	-0.098*	-1.7
Membership in credit and savings groups	0.023	1.62	0.004	0.22	0.005	0.28
Membership in cooperatives	0.035**	2.31	0.055***	2.86	0.047***	2.81
Poor soil	-0.018	-0.39	-0.001	-0.02	-0.027	-0.51
Medium soil	-0.015	-0.88	-0.055**	-2.26	-0.021	-1.08
Men managed plots	-0.070**	-2.51	-0.018	-0.45	-0.058*	-1.86
Men and women managed plots	-0.024	-0.84	0.021	0.5	-0.010	-0.32
Poundability	0.000	-0.01	-0.010	-1.08	0.004	0.54
Quality of garri	0.022*	1.74	0.009	0.48	0.006	0.4
Quality of fufu	-0.001	-0.07	0.024	1.53	0.013	1.07
Ease in loading	-0.010	-1.49	-0.016*	-1.71	-0.016**	-2.08
Starch content	0.012*	1.78	-0.003	-0.35	0.015*	1.91
Tapioca taste	0.002	0.22	-0.008	-0.78	0.004	0.44
Market demand	0.007	0.66	0.023	1.45	0.022*	1.66
Root size	-0.013	-0.9	-0.002	-0.12	0.005	0.27
White color	-0.005	-0.71	-0.003	-0.28	-0.009	-1.1
Stores well	-0.038***	-4.21	-0.034***	-2.84	-0.033***	-3.21
Root yield	0.001	0.04	0.005	0.19	-0.028	-1.39
Pest resistance	0.010	1.22	0.012	1.1	0.006	0.65
Stem yield	-0.018**	-2.37	-0.027***	-2.65	-0.024***	-2.87
Early maturity	0.011	0.95	0.032**	2.08	0.020	1.46
Yellow roots	-0.013**	-2.01	-0.005	-0.53	-0.013*	-1.82
N	4468		4468		4468	

*Regional dummies are included but not reported. ***, ** and * refers to significant at 1%, 5% and 10% level.*

Who are the farmers that correctly identified improved varieties?

In this section, we ask the question of who are the farmers who correctly reported improved varieties. We measure correct classification by a dummy variable which takes a value of one if there is an agreement between DNA based varietal identification and self-reported results and zero otherwise.

Importantly, three sources of heterogeneity could largely explain the difference in correct identifications. These are difference in education, access to information and location. Irrespective of the scenarios considered, households who correctly reported improved varieties are more likely to be more educated. In all the scenarios considered, education has a positive and significant effect on the probability of correctly identifying improved varieties that the farmers grow in their plots. In terms of our second source of heterogeneity, access to information, we found that households who correctly reported improved varieties have access to extension, own mobile phones and are a membership of cooperatives. As such households who have access to more structured information sources are more likely to be informed about the type of variety they grow in their plots. Our final source of heterogeneity is location. We found that farmers located in the northern, south-western and south-south part of Nigeria are more likely to correctly report the type of variety (in terms of improved and landrace) compared to those located in the Southeastern part of Nigeria.

Table 32. Determinants of correct classification.

Variable	Scenario 1	Scenario 2
Household Size	0.0381***	0.0417***
	(0.0142)	(0.0143)
Education	0.0160**	0.0109*
	(0.00649)	(0.00649)
Marital status	-0.0654	-0.112
	(0.129)	(0.129)
Age	-0.00671	0.00618
	(0.0144)	(0.0143)
age2	0.000103	-2.94e-05
	(0.000132)	(0.000132)
Sex	0.0470	0.134
	(0.133)	(0.133)
TLU	-4.27e-05	3.62e-05
	(0.00303)	(0.00308)
Access to credit	0.125*	0.0713
	(0.0673)	(0.0672)
Access to extension	0.253***	0.275***
	(0.0689)	(0.0688)
Distance from market	-0.00961*	-0.00770
	(0.00543)	(0.00509)
Availability of cassava stems	0.251***	0.114
	(0.0725)	(0.0720)
Radio ownership	0.0193	0.0440
	(0.0866)	(0.0863)
Mobile phone ownership	0.598***	0.630***
	(0.169)	(0.163)
Television ownership	-0.0803	-0.106
	(0.0773)	(0.0773)
Membership in credit and saving groups	0.125*	0.108
	(0.0655)	(0.0659)
Membership in cooperatives	0.207***	0.230***
	(0.0704)	(0.0712)
Value of farm assets	-0.00141**	-0.00130*
	(0.000679)	(0.000680)
North	0.548***	0.615***
	(0.111)	(0.110)
Southwest	1.039***	1.086***
	(0.109)	(0.108)
South-South	0.884***	0.793***
	(0.101)	(0.100)
N	2024	2024

*Robust standard errors are reported in brackets. ***, **, and * refers to significant at 1%, 5%, and 10%, respectively.*

Component 3: GPS-assisted area measurement

Take-home message:

- GPS-based measurements of area and measurements based on self-reporting produce different results. In particular, farmers overestimate the size of small farms and underestimate the size of large farms.
- The measurement error (the discrepancy between GPS and self-report area) is normally distributed.

Why GPS-assisted area measurement?

In the mainstream development economics literature the inverse relationship between farm size and productivity is taken as a stylized fact (Collier 1983; Van Zyl et al. 1995; Barrett 1996; Barrett et al. 2010; Carletto et al. 2013). However, such relationships are mostly based on self-reported area measurement which is prone to measurement errors. In many household surveys, farm sizes are measured by simply asking farmers to estimate the area of their farm plots. While this approach is simple and inexpensive (in terms of money and survey time), it may produce imprecise values for many reasons: first, most farmers have limited understanding of the purposes of surveys. As such they may report wrong values, if they perceive that the information collected by enumerators is to be used as a basis for taking away their land as part of a large land re-distribution policy (Holden and Fisher 2013). This is the situation in the context of many developing countries as land tenure systems are quite fragile or non-existent. Secondly, due to the low level of formal education and underdeveloped land markets, farmers are often unfamiliar with standard units of area measurement units (Dillon *et al*. 2016). Even when local units of measurement are used, conversion to standard units is also prone to errors. Therefore, estimates of farm size based on self-reporting could be considered inaccurate. However, these errors can be consequential as ex-post assessment of the benefits of adoption heavily relies on the accurate measurement of land size and production values.

In an effort to measure the level of measurement errors, the CMS project introduced Global Position System (GPS)-based area measurement in Nigeria. This serves to measure the magnitude of self-reported bias as well as to undertake a credible causal analysis of adoption impacts on poverty and food security outcomes. This is particularly important as the discrepancy between self-reported and GPS-based area values may become substantial and in some cases it may vary by farm size or level of formal education. The effort by CMS to use GPS-based area measurement will therefore contribute to this much debated relationship between farm size and productivity. In Table 33, we report summary statistics for the cassava area for which we have information based on GPS and farmers self-perception. For now, we take the GPS measure as the benchmark. The level of bias (under/over-reporting in absolute terms) is then presented by comparing cassava area as measured by GPS and by self-report. Over-estimating refers a situation where the area reported by the farmer is larger than the GPS measurement. Average cassava area based on farmers' self-reported value and as measured by GPS produce different estimates. The mean difference between the two methods is about 0.03 ha and this difference is statistically significant at 1% level. In addition, the distribution of cassava area across the different States suggests that such underestimations are prevalent in all States, except in the northern part of the country. Results show that even though the average difference is statistically significant, the magnitude of overestimation is very low.

Table 33. GPS vs self-reported.

Level	GPS Mean(ha)	Self-reported mean(ha)	Mean diff
Pooled	0.34	0.31	0.03***
North	0.461	0.457	0.004
Southwest	0.469	0.457	0.012*
Southeast	0.15	0.144	0.006**
South-South	0.41	0.31	0.1***

Table 34 summarizes the size of the bias by cassava land deciles. As before, the discrepancy is measured as the absolute difference between the GPS measure and self-reported area values. The result suggests that the magnitude of negative bias (overestimation) appears to increase from the bottom to the largest decile except for the last four deciles.

On average, the absolute bias in the bottom decile is about 0.013 ha. In the last decile, the magnitude of the bias is reversed (underestimated) by about 0.27 ha. Our data suggest that there is a pattern in the direction of the bias whereby smaller farmers generally over-report their land compared with larger farmers. Figure 17 also shows the distribution of cassava plot areas based on GPS and self-report over all respondents. As the distribution graph can easily show, there is a strong correspondence between the GPS and self-reported area measures. The distribution of self-reported area measurement largely mimics the distribution of GPS-based area values although with a clear difference at the lower (left tail) and top (right tail) end of the distribution. In addition, the distribution shows the same tendency where farmers overestimate the size of small farms and underestimate the size of large farms.

Picture 11. Enumerator measuring cassava fields during CMS survey using hand-held GPS.

Table 34: Distribution of measurement errors by land decile.

Quartiles	Mean area using GPS	Mean area self-reported	Mean diff
1	0.017	0.03	-0.013***
2	0.045	0.074	-0.029***
3	0.074	0.08	-0.005***
4	0.108	0.137	-0.029***
5	0.152	0.18	-0.025***
6	0.21	0.22	-0.01
7	0.294	0.28	0.016**
8	0.42	0.37	0.05***
9	0.67	0.60	0.07***
10	1.42	1.15	0.27***

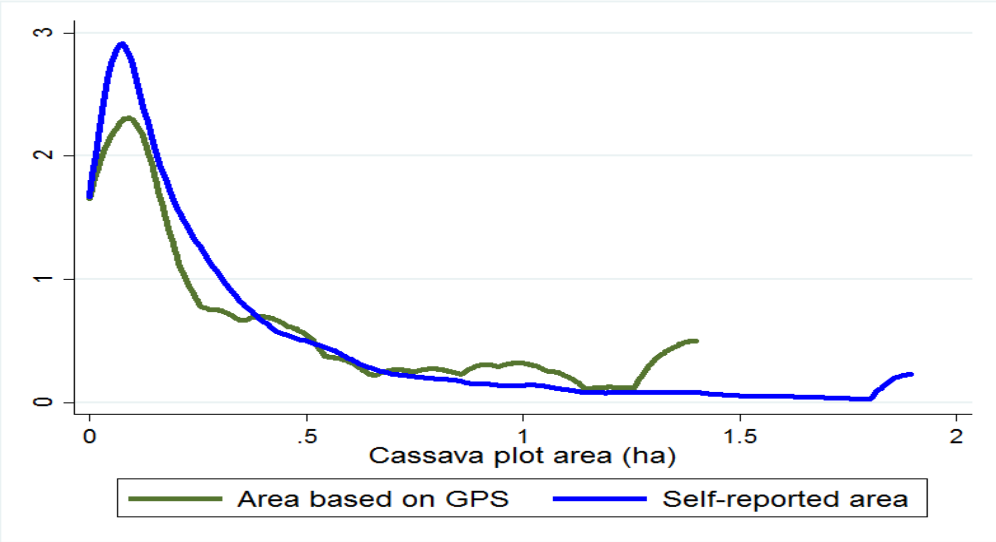


Figure 17. Distribution of GPS and self-reported area

Moreover as shown in Fig 18, the two distributions exhibit a very high correlation
Given our reported results, the obvious question is whether we should use GPS in all household surveys or continue to rely on self-reported values? Our results clearly show that measurement error in land size has the potential to distort important and policy-relevant econometric relationships especially when land size is included (either directly or indirectly) in the set of independent variables. However, as the distribution in Figure 19, shows, the errors are actually quite noisy but normally distributed.

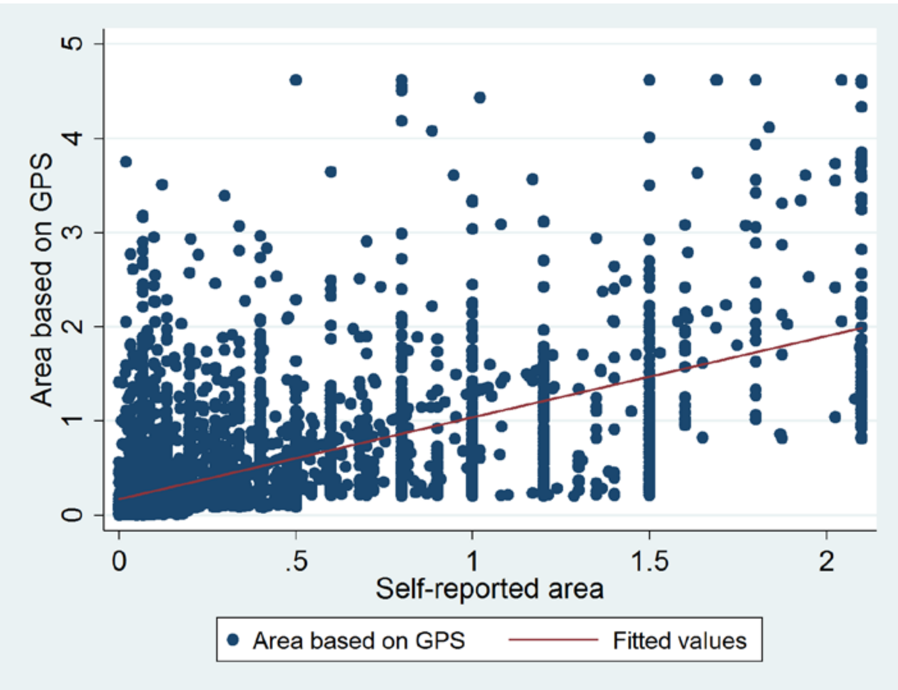


Figure 18. Correlation between GPS-based plot area and self-reported plot area

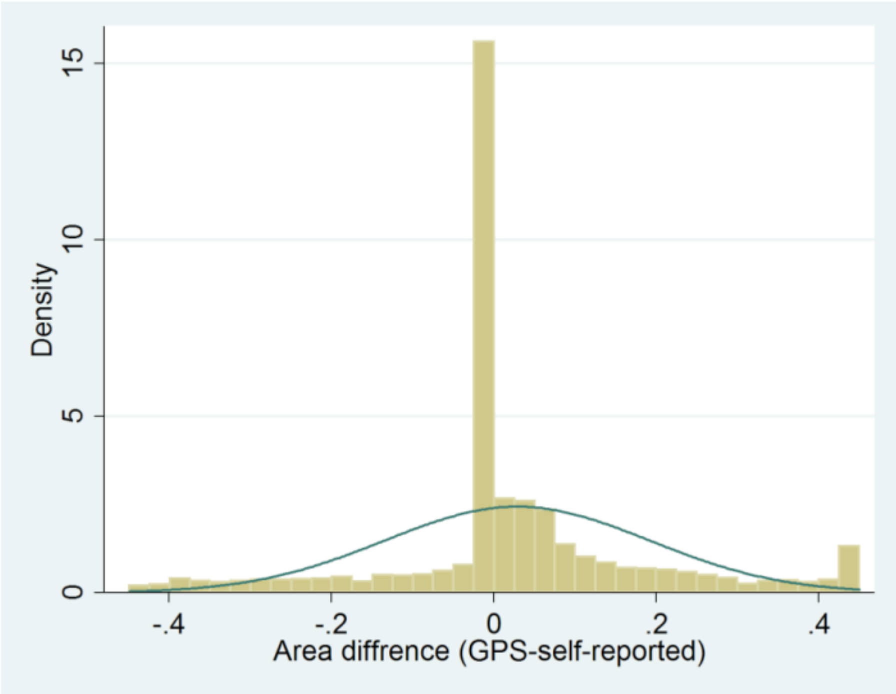


Figure 19. Distribution of measurement biases.

Moreover, while comparing adopters and non-adopters, we did not find any statistical difference in terms of the magnitude of reported bias. As such, causal inference in adoption studies can still be robust even in the presence of land size measurment errors. However, these results are often context-based and have to be interpreted with caution. Generally, if precision is central, then GPS-based measurements are important. However, by using GPS- based area measurement, we may introduce another measurement error when production volumes are collected, based on farmers' self-perception instead of actual measurements. As such, it is important either to measure production and area using physical measurements or rely on self-reported values. Note that measurement errors will not introduce bias when the variable measured with error is used as an outcome indicator (dependent variable).

Component 4: Gender-differentiated end-user surveys on varietal and trait preferences

This section presents the results of gender-differentiated varietal trait preferences based on village-level focus group discussions and using a sub-sample of the CMS data where we collected detailed preference related data from 30% of the spouses.

What farmers like about the varieties they grow

Research question: *Why are farmers growing the major varieties identified in the CMS survey?*
Nigerian farmers like the cassava varieties they have because they are:

- High yielding: many big roots
- Early maturing
- Durable (store well underground for at least two years).
- Also: tolerate poor soils, cattle resistant, not watery, and drought-resistant (in the North)

A few traits are optional or variable, e.g., some farmers like more starch, some like less. There is a demand for both yellow and white roots. And even though farmers express a demand for precocious varieties, farm communities may also need some that are late maturing. But, in general, trait preferences are straightforward (e.g., every variety can be high yielding, whether targeting women or men, or any region). For processing traits, season and age are more important than variety. Agronomic traits are, in general, more important than processing qualities. Women demand cassava that is easy to peel, although this need could possibly be met with another research solution, not by plant breeding.

Each community grows some four to eight varieties; about two of these are usually so important that every household grows them. In each region, a handful of varieties predominates, and is grown in three or four communities. The others are old varieties on their way out, new ones that have reached only a few communities, or the odd variety that a traveler brought back home some years ago. Few varieties are grown in more than one region. Men and women tend to grow the same varieties. Some of the improved varieties arrived fairly recently, including some from ADPs that are no longer able to distribute planting material.

What farmers like in specific varieties

In general, the improved varieties have the qualities that farmers like, and are gradually squeezing out the local varieties. Improved varieties are becoming the standards because they are early maturing and high yielding. Farmers want all varieties to be:

- High yielding (many big roots).
- Early maturing (ideally reaching maturity within six or seven months).
- Able to be stored underground for a long time (at least two years, i.e., one year or more after maturity).

These are prerequisites, a short list of specific agronomic traits that all varieties must have. If some of the varieties that store well in the ground are late maturing, that might be an acceptable trade-off. While every household needs some cassava that is early maturing (to break the hungry season) it is also important to have cassava all year round, like a bank account from which people can steadily withdraw for food and to make money. There are also some traits that are required only in some varieties, or in some places. For example, some communities prefer shorter varieties which are less likely to topple over in the wind. Other villages are pleased with their tall cassava varieties. Like the women, men do mention processing and culinary traits (e.g., going well with soup).

“Poundable” means “not toxic,” and does not refer to the texture of the root. Poundable cassava can be boiled, pounded, and eaten with no further processing, because the variety is sweet, or nontoxic. It is somewhat odd that farmers mention “poundability” so often. They even express a preference for poundable varieties in communities where people always process their cassava and have not eaten it boiled for a generation. Many adult Nigerians have never eaten freshly boiled cassava in their entire lives, yet for some reason they pay attention to this trait. It may just be worth remembering which varieties are toxic and which are not, just in case a person ever needs to eat some fresh (boiled) cassava. In some regions like the South-South and Southeast, people make a cassava food called *abacha*, which is not fermented, so it must be made from nontoxic varieties. The women mention processing traits more often than the men (e.g., roots being easy to peel), because women process most of the cassava. When cassava is difficult to peel, the women end up cutting off some of the flesh with the peel. When cassava is hard to peel it not only lowers the efficiency of women’s labor but also the economic yield.. Women also notice the agronomic traits of cassava varieties[9].

In the Southwest: There are a few quite specific reasons for some varieties to be preferred For example, the women of Ajagbale village in the Southwest like black cassava because it makes good shade when intercropped with cocoa. The young cocoa plants can be planted under the cassava plants. (No doubt the farmers then harvest the cassava and replace them with a definitive shade plant.)

In the North: When describing what they liked about each variety, women and men in the North gave more or less equal attention to agronomic and processing traits. Some people in the North mentioned cassava for making dried cassava (*alibo*), which can be stored for a long time. When the root is nontoxic, the cassava leaves can be cooked and eaten as a sauce. Some cassava varieties are so nontoxic that they are safe to eat raw.

In the South-South: People mention that some varieties can be intercropped with oil palm, which is the most important cash crop in this region. Some people mention late maturing as a desirable quality. While “easy to peel” is the most important trait for women (see Sections 2 and 3.2), ease of peeling rarely shows up when varieties are described: probably because most varieties are equally difficult to peel (but see AbegheTighe, on Table 11). In the South-South some people make *abacha*, which is not fermented, and so it must be made from nontoxic varieties.

In the Southeast: For some reason, the farmers mentioned processing less than in other regions, although they did single out a variety which is easy to peel (*Nwanyi Ocha*), which is important but unusual. People make *abacha* in the Southeast, and so will need some nontoxic varieties. Villagers like several varieties because they are low in fiber and so good for making *abacha*. Unlike the Southwest and the North which praised short, wind-proof varieties, in the Southeast farmers value some of the tall varieties (e.g., Police Cassava), some of which are tall enough to shade people as they rest in the field. Farmers in this region like some varieties just because they have an unmistakable leaf and stem color and so are good for marking field edges. This suggests that agriculture is so intensive that people are worried about their neighbors encroaching on a field, and the land is too valuable to be wasted by putting up a fence, so a living, edible boundary marker is ideal.

Improvements farmers want in the varieties they grow

In all four regions, even some popular varieties came in for constructive criticism. Again, men do mention processing qualities. Women give importance to agronomic traits, not just processing. There is a certain consistency for preferred traits, across regions and genders.

[9]Detailed preference heterogeneity at the regional level can be found in the supplementary material.

Preferred agronomic traits: high yielding, fast maturing, store well underground, nontoxic.

Preferred processing traits: nontoxic, not watery, low in fiber, white for making fufu. *Gari* and fufu should be moldable, meaning that it can be rolled into a ball before it is swallowed. Women and men seldom mention ease of peeling, which suggests that they realize that most varieties are equally challenging to peel.

In the Southwest: Farmers, besides wanting improved varieties to be high yielding, early maturing, and to last for a long time underground, also want varieties that suppress weeds. Women want varieties that are easy to peel. Women had some criticisms of the improved agricultural varieties. Farmers want short varieties that can withstand the wind. Some varieties are susceptible to grasscutters (a large rodent, the greater cane rat). Some traits are also contradictory: for example, farmers would like a variety that the Fulani cattle do not eat, but dislike extremely toxic varieties, which can poison their own goats.

In the North: Resistance to drought is mentioned, once. Some varieties are too poisonous. Farmers want fast maturing varieties that can stay in the soil for two years or more.

In the South-South: The men want varieties that produce well in infertile soil, and need little weeding. Both genders want cassava that matures quickly and resists rot (so it can stay in the soil for a long time). Some varieties are too bitter.

In the Southeast: women and men are more concerned with agronomic traits, for some reason, and had little to say about processing (Table 16). The women make a sophisticated observation: varieties with broad leaves block out the sunlight, keeping the soil moist and promoting rot. Villagers want varieties that store well in the soil. Some varieties are susceptible to attack by grasscutters (cane rats), and Fulani cattle. The men make a rare observation about peeling for one variety (Ekpe Calabar). The women would like to plant more Vitamin A cassava.

General preferences for cassava traits

Research question: *Are there any general preference traits farmers are looking for?* In the previous section, the farmers were asked what varieties they planted and what were the advantages and shortcomings of each one. We saw that there are great regional differences in varieties planted, but only slight differences in the perceived advantages and disadvantages of these varieties (e.g. in all regions women and men see early maturity as a good trait). When the team asked what traits the farmers preferred in cassava in general (not variety by variety), the farmers mentioned a few new desired traits. All the women want cassava that is easy to peel. Women and men want cassava that is high yielding. Most place a high value on cassava that is early maturing and stores well underground. Table 35 presents the main preferred general preference.

Table 35. Trait preferences by women and men in Nigeria.

Community	Women	Men
Southwest	Easy to peel, processing high yielding, early maturing	High yielding, early maturing, stores underground, controls weeds, ready market
North	Easy to peel, high yielding, nontoxic, stores underground, processing	Early maturing, Insect resistant High yielding, Access to market
South-South	Easy to peel, high yielding stores underground, processing	High yielding, stores underground Tolerates poor soils, Early maturing
Southeast	Early maturing, Easy to peel, stores well underground Big roots (high yielding)	Fast maturing, high yielding Less starch, drought resistant

Gender differences

Research question: *Are there differences among men and women farmers in varietal preference in each study region and across regions?* Some men want more market for fresh roots but women already have a thriving market for their products. We have also seen that women pay more attention in general to processing. In this section we will see that the common varieties are popular with men and women. Gender differences (varieties grown only by women or only by men) show up only with the uncommon varieties, and this may be due to sampling error. (e.g. if a variety is grown by only two people in a village, a man and a woman, the woman may attend the FGD, but not the man, or vice versa). In each village, the most frequently planted varieties are the same for men and women. In the previous section, we reported that women and men prefer varieties that are fast maturing and high yielding, that store well in the ground. For processing, farmers want nontoxic varieties that are low in fiber and white when made into fufu. They want varieties that are low in moisture (although they know that all the roots are moister in the rainy season and drier in the dry season).

Box: Why women want varieties that store well in the ground

Some women cultivate several small plots of land and like to harvest them a few plants at a time. Varieties that keep well in the ground can be harvested, processed and sold any time women need money for expenses such as school fees, clothes for the family, medical bills or food items that the household does not produce. The cassava stored in the ground is collateral to borrow money from people ("***Loan me a bit of money and I will repay you when I harvest my cassava and make*** garri"). When there is cassava in the garden, there is food in the house all year round.

Both men and women paid attention to agronomic and processing traits, and saw more or less the same advantages with each variety. For women, processing traits are crucial, especially ease of peeling (which seems to be a real problem for them, especially since grating is mechanized). Most varieties are similar for toasting and grating: the difference between roots depends on their age (old ones are tougher) and season (roots are more moist in the rainy season). For the men, agronomic traits are more important than processing qualities. Some men, especially in the Southwest and the North, would like to be able to market cassava as fresh roots without having to process it at all. In some communities the women are more involved with cassava and know more about it than the men. The men rarely mention ease of peeling (which is the top preferred trait for women) but do have an idea that some processing traits are important (such as color). Women are keenly aware of processing and agronomic traits.

Processing traits are more important for women than for men

In all four regions, the men pay attention to processing but rank it as less important than agronomy (e.g., yield). In the Southwest, for example, men rank processing traits sixth (high dry matter content), eighth (stable color), ninth (swelling ability), and eleventh (lots of starch). Culinary traits such as Vitamin A and edible leaves rank even lower. The other regions are similar: yield, maturing quickly and good storage underground are at the top of the list. Occasionally even men mention that some varieties produce leaves which are good to eat as vegetables. Women, on the other hand, rank processing traits as the most important: being easy to peel is in first place in three regions and in the second place in the other. Women want varieties that are nontoxic ("poundable" even though rural Nigerians no longer pound and eat much cassava without fermenting it). Color is also important: white for fufu *(akpu*) and *abacha* and sometimes for garri, while yellow is the ideal color for most garri.

Dewatering fast, being moldable, swelling when mixed with water, and retaining its color in processing all rank highly.

However the women realize that most of these traits depend on the season of the year or processing method. For example, cassava is easier to peel when it is young, and in the rainy season. Cassava mash is drier and hence easier to drain of water in the dry season. Moldable fufu and garri *(eba,* actually) depend on being properly fermented and processed. The women also want the same key agronomic traits that men appreciate (high yielding, early maturing, and storing for a long time underground), but these are less important than simply being able to peel a large mound of roots quickly so they can get on with their many other daily tasks.

Men and markets: Some men demand ways of marketing cassava fresh, without having to process it. Women are far less interested in selling raw cassava because they do the processing. Selling the cassava as garri or fufu allows women to earn some money, an opportunity that they would lose if the roots were sold raw.

Storage pests: One way to store cassava is to dry the roots as chips, or *alibo*. There is, however, a storage pest in the North, a beetle (possibly *Prostephanus truncatus*) which bores galleries (neat round tunnels) through the dried chips, quickly destroying the whole store.

Poundable means nontoxic: In the first few villages in the Southwest, farmers told the team that certain varieties were good because the roots could be boiled and pounded, like yam. It turns out that most people haven't eaten pounded cassava since they were children. Some have never eaten it. Now people have yam to pound, and cassava is most often made into garri. If you ask people what is good about a variety, they may respond with a memory of mom's cooking, but not all varietal traits really matter.

An advantage for Vitamin A cassava: Many consumers prefer a buttery-colored garri. So far, farmers are interested in the Vitamin A cassava because it is yellow. This may allow village-level processors to make garri without adding palm oil, thus saving an expense. Oil affects the taste and texture as well. Still, most Nigerians prefer yellow garri, so a yellow cassava is not out of place.

Trait preferences of agro-industry: In the future, if more food manufacturers begin to use cassava as a raw material, these industrial processors will express demands for specific traits, such as higher starch content.

Access to planting material

Research question: *Why are farmers growing only local varieties and not using any improved varieties?* Farmers do not use more improved varieties because of a lack of access to the seeds. Most of the farmers, women and men, are growing improved varieties. When they get a new improved variety, farmers plant it, and evaluate it, and then usually plant it on a larger scale. Improved varieties tend to replace local varieties. A village may grow a variety for 30 to 50 years. Many improved varieties arrived in the 1980s or even 1970s, often through the ADP of the state government. Some villages have received planting material since the 1980s while others have not. Some have had sporadic access to new varieties from faith-based organizations or NGOs. Since about 2010 or 2012 there has been some distribution of improved varieties, including Vitamin A cassava. Some of the ADPs are still able to distribute planting material. Adoption of improved varieties can be quite rapid. People in Ajagbale village in the Southwest acquired a white variety, perhaps as recently as 2014, from the ADP and IITA. All the farmers are now trying it so they can sell the cassava to a starch processor.

Table 36. Access to cassava seed by women and men in Nigeria.

Community	Women	Men
Southwest	Mostly they share with each other. Some buy	They share and never buy
North	They share. Only one village knew of buying	The men share. In three villages a few buy stems
South-South	They share and buy	Same as for women
Southeast	They buy from farmers and in the market and sell among themselves. They also share	Same as for women

Men and women in all four regions rely mainly on sharing with each other and on purchase when they cannot supply enough seeds from their own farms (Table 36). Many farmers have little or no direct access to improved varieties from the formal sector, and improved varieties are not trickling very fast through the informal sector. New varieties are best distributed early in the rainy season. Some people recently received Vitamin A varieties in the dry season and the plants died.

In the Southwest: Many people received early improved varieties, about the 1980s, and started to receive improved varieties again only after 2010. In some communities (Erin Oke and Ilesa) people were not able to properly track recent improved varieties. Sometimes, when several improved varieties are introduced at once, people get them mixed up.

In the North: Northern farmers have had sporadic access to improved varieties, especially since about 2000. They often search out new varieties when they travel, or source varieties from neighbors. Farmers have a real interest in trying new varieties. Improved varieties can spread spontaneously, hand-to-hand, although this is not necessarily the quickest or most systematic way for a variety to be distributed.

In the South-South*:* There has been some recent access to improved varieties from NGOs and faith-based organizations. Farmers in some villages have had access to some early improved varieties since the 1980s or late 1970s. During the 2000s, access increased, especially through the ADP. All five villages have had at least some access, especially to a recent popular variety called Give Me Chance.

In the Southeast: People seem to have less access to improved varieties than in other regions. They don't always know if a new variety is improved or not but they do grow some new varieties, including improved ones. They buy more stems than in other regions, often going to the weekly market and buying mixed bundles of unknown varieties. Men don't grow much cassava; the women do, but the men are now trying to get into the crop. Cassava has been in the custody of women, so they know more about it, and may have more varieties. The village of Oraifite was an unusual case. Many of the local people were university educated and said that they had been receiving IITA varieties since the 1980s. Since then, the villagers have grown these varieties, but their favorite was still a local variety, NduKaN'ala, which was high yielding and early maturing.

Access to cassava seeds

Buying cassava stems is not common, but does occur. Generally each household produces its own cassava stems. If people need more stems they can usually get them for free, from a neighbor who has just harvested a field. If people have lost their cassava to drought or if they want to expand their fields they may need more stems, but usually can get them for free. Most people never buy cassava stems. However, even in communities where some people buy seed, the most common way to acquire seed is to use stems from one's own garden or to ask permission to collect stems from the field of neighbors who have just harvested. Getting stems from neighboring farmers lowers transaction and transportation costs. People who grow lots of cassava have lots of stems, and less cause to buy planting material. Men in the Southwest and in the North, for example, grow much cassava and have less reason to buy seed. Where palm oil or another crop is a major cash crop, farmers grow less cassava, run out of seed sometimes, and have to buy it. Where people do buy seed, in the weekly market, the varieties can be all mixed up.

Men and women access seed differently: At the Cassava Growers' Association in Ilesa, Osun State, the women explained that since they grow less cassava than the men (and process more), they are likely to run out of stems sooner. The men can leave a part of their field unharvested, saving the stems for planting time. The women may not always be able to do so, but the women can "buy a cassava farm," meaning that they buy the standing crop in the field, harvest it for the roots, but also keep the stems. The women said "It's our trick." The men may not even realize that they are selling seed, when the women buy the standing crop. The men in those villages may have larger fields, and so are less likely to run out of stems.

In the Southwest: Men and women prefer to produce all of their own stems. When they cannot produce enough stems on their own farms, women and men first try to borrow stems from neighbors. There was a gender difference in some communities, where men said they never, ever bought stems, but women said they often did. Sometimes the women "bought a farm" (i.e. bought a crop of cassava standing in the field), often from a man. The women harvested the field, kept the roots, and planted the stems. The farmer who sold the crop may not have even been aware that he was selling seed.

In the North: By far, the most common way to get seed is to grow it by one's self, or to get it for free from neighboring farmers, usually those who have just harvested their cassava and have stems left on the ground. Some people buy cassava stems "along the road," i.e. from the little roadside stalls. In the North, some people have been forced to buy stems because herders have allowed their cattle to graze on and destroy cassava fields.

In the South-South: People produce their own stems and exchange planting material with neighbors, but there is also more buying and selling of stems there than in the Southwest or the North.

In the Southeast*:* People share stems, but they also buy them from each other and from the market. They buy stems more than in other regions. This may be because cassava fields are smaller, so it is easier to run out of planting material by the start of the rainy season. In some communities e.g. Amugo in Enugu in the Southeast, improved varieties flow through the system, even if there is little contact with the formal seed sector.

Disadoption of cassava varieties

Research question: *Are there examples of disadoption and what are the main reasons for such behavior?* Disadoption is mostly of old varieties that were replaced by improved varieties They were disadopted because they were low yielding, late maturing, did not control weeds, fell over in the wind, were too toxic, etc. Improved varieties are not always the favorites but are usually high on the list of popular varieties for men and women in each region. There is little or no disadoption of improved varieties. Seeds of a variety can be lost. If people get a new variety and everyone plants it, they may lose the planting material for an old variety without really meaning to do so or without realizing the situation until it is too late. There's no place to store seeds, so they collect seeds from neighbors. If they harvest all their cassava in the dry season, they have to get seeds from somewhere else when the rains start. Farmers are gradually abandoning local varieties but most villages have never abandoned an improved variety. In future, as new improved varieties are released, they may begin to replace the older improved varieties. The local varieties are abandoned because they are late maturing, poor yielding, begin to rot quickly underground, or do not control weeds well.

In the Southwest: In Ajagbale, for example, farmers are abandoning some local varieties, usually for explicit, agronomic reasons. *Ege pupa* (red cassava) is low yielding and late maturing. This is also one of the few communities where people say they are discontinuing TME-419. Although some people are still planting it because of its high yield, the others say it falls over in the wind. It may also be facing stiff competition from the starchy white cassava. Note that many of the abandoned varieties are mentioned only in one community, and may have been rare to begin with. Few improved varieties are abandoned.

In the North: Farmers have probably disadopted only local varieties. The abandoned varieties were too toxic, too slow to mature, or yielded too little.

In the South-South: Most of the disadopted varieties were local: a few were old improved varieties, which farmers had grown for years and were starting to abandon in favor of newer ones. Varieties are discarded for being poisonous, too watery during processing, or because they rot easily (store poorly underground) or are poor yielding.

In the Southeast: Some local varieties were replaced by improved ones. Varieties are abandoned for being low yielding, or late maturing. Some were also too susceptible to attacks by grasscutters, monkeys, or other vertebrates.

Does gender based preference heterogeneity affect adoption decision

With a focus on improved varieties, this section answers two important empirical questions: *(1) Are there any systematic differences in preference trait heterogeneity between the spouse and the HHH? (2) If so, do these differences affect adoption decisions?* Answering these questions will help breeders, donors and policymakers to provide recommendations that take into account gender perspectives. The CMS project collected gender-disaggregated data on trait preferences from a sub-sample of randomly selected villages[10]. In particular, we collected gender preference data from 30% of the spouses of the households interviewed in the main survey. Although there has been a plethora of empirical research on the determinants of technology adoption and, to a lesser extent, on the role of preference heterogeneity, research is rather scant on how and why gender-based preference heterogeneity may affect decisions on adoption and disadoption. Preference trait heterogeneity between the spouse and the household head are important for adoption at the farm level. In particular, when adoption choices are not made based on collective decision, as predicted by a unitary household bargaining power model, trait preference between the spouse and the household head will play a crucial role for the adoption of improved varieties. Failure to account for this source of heterogeneity may therefore lead to wrong conclusions about adoption decisions at the farm level.

Table 37. Descriptive statistics for gender based preference

Variable	Household head	Spouses	Mean diff
Adoption	0.63	0.55	0.08***
Household Size	4.54	4.27	0.27***
Education (years of education)	9.22	8.07	1.15***
Age	50.7	40.65	10.05***
TLU	0.64	0.897	-0.18
Access to credit (1=yes)	0.68	0.50	0.18***
Access to extension (1=yes)	0.56	0.39	0.17***
Distance from market (km)	3.13	3.13	0.000
Radio ownership (1= yes)	0.86	0.84	0.02*
Mobile phone ownership (1= yes)	0.96	0.96	0.000
Television ownership (1= yes)	0.76	0.71	0.05***
Membership to credit and saving groups (1=yes)	0.34	0.43	-0.09***
Membership to cooperatives	0.25	029	-0.04
Poundability	2.9	3.07	-0.13***
Quality of Garri	3.7	3.7	0.000
Starch content	2.57	2.83	-0.25***
Quality fufu	3.65	3.66	-0.001
Tapioca taste	2.3	2.4	-0.01*
Root Yield	3.77	3.75	-0.02
Early maturity	3.67	3.7	-0.03
Easy to load	2.23	2.66	-0.43***
Market demand	3.64	3.64	0.000
White color	3.12	3.53	-0.41***
Stores well	3.54	3.67	0.01
Pest and disease resistant	3.27	3.12	0.15***
Stem yield	2.99	3.11	-0.12***
Yellow root color	2.03	2.4	-0.38***
N	758	758	758

[10]The results of this section will further be complemented by qualitative case studies in selected locations on gender differences and gendered decision-making to capture the role of gender in adoption and to get a deeper understanding of adoption pathways, varietal preferences, and traits. Other end-users include a large number of commercial processors (medium-to-large scale or fully mechanized processors of garri, flour, instant fufu, native starch, ethanol, feed, etc.,) whose preferences differ from those of households would need to be captured in this survey as well.

Herein, we check the consistency of the results reported in the previous section using our sub-sample of gender-disaggregated data. Failure to account for this source of heterogeneity may therefore lead to wrong conclusions about adoption decisions at the farm level. Herein, we check the consistency of the results reported in the previous section using our sub-sample gender disaggregated data. Using our gender disaggregated data, only 55% of the spouses have reported to have planted improved cassava varieties in 2015. However, the response from the household head shows adoption rate of 63%. The mean difference in the reported adoption rate from the household head and spouse is statistically significant at 1%. This difference may arise as some plots are solely managed by either the spouse or the household head. As such, for plots managed by the spouse, cassava-related questions should be directly addressed to the spouse instead of the household head. In addition to reported adoption rates, we found significant differences between the spouse and the household head for other variables of interest such as access to credit and access to extension. The mean differences in extension and credit access are also statistically significant at 1% significance level. Data from spouses show that only 50% of the spouses have reported to having access to credit while about 68% of the household heads reported to having credit access. Note that, credit access is constructed using a village threshold level of 25%. Similarly, the descriptive statistics on spouse and household heads show a striking difference. For instance, only 39% of the spouses have access to extension. However, about 56% of the household heads have reported to have extension access. Similar to credit, extension access is constructed based on a 25% threshold level. The mean difference in extension access between the spouses and the household head is also statistically significant at 1%.

Coming to our main variable of interest-preference heterogeneity in improved cassava varieties between the spouse and household head, we found striking results. Generally, spouses tend to attach higher values to cassava traits compared to household heads (where about 98% are men). In particular, we found statistically significant differences between the spouses and household heads for the following preferences: poundability, starch content, tapioca taste, ease of load, white color, stem yield, yellow color, and resistance to pest and disease. With the exception of resistance to pest and disease, the weights attached for these traits are always higher among spouses. For traits such as garri taste, root yield, market demand and storage, we did not find any significant difference between the spouses and the household head. The traits, ease of loading and poundability, have been considered as unimportant by household heads.

However, spouses tend to consider this trait as important, suggesting that labor-intensive activities are mainly done by spouses. The observed differences in the adoption rate between the spouse and the household head cannot simply be attributed to differences in trait preference by looking at mean differences between the spouses and the household heads. In particular, these changes are only indicative of correlations and cannot be used to make causal inferences regarding the impacts of trait preference without controlling for confounding factors. In the next section, we implemented an econometric model to empirically estimate the effect of preference heterogeneity between spouses and household heads on the adoption and disadoption decision.

Econometric analysis on gender-disaggregated data

Our gender disaggregated probit regression results on spouse and household level data reveal that preference trait heterogeneity between the spouse and the household head affects adoption decision differently. We found that important traits such as early maturity, quality of fufu and garri, root yield, and storability have statistically significant effects on the adoption decision of both spouses and household heads. However, traits such as stem yield, poundability, and pest and disease resistance, affects adoption decisions of only spouses. Finally, the trait, ease of loading, affects the adoption decision of only household heads. In other words, by asking the spouse and the household head the same question, determining factors of adoption can be different. For some traits such as tapioca taste, effects were positive and significant when using responses from spouses but negative and significant

Table 38. Probit results on gender based trait preference.

	Household head		Spouse	
Variable	Parameter estimates	Marginal effects	Parameter estimates	Marginal effects
Household Size	0.0801***	0.026***	0.0238	0.0086
	(0.025)	(0.025)	(0.026)	(0.026)
Education	-0.0170	-0.005	-0.0042	-0.0015
	(0.011)	(0.011)	(0.012)	(0.012)
Age	-0.0005	-0.00016	0.0050	0.0018
	(0.004)	(0.004)	(0.005)	(0.005)
TLU	-0.0204***	-0.0066***	0.0014	0.0005
	(0.004)	(0.004)	(0.006)	(0.006)
Access to credit	0.2073*	0.067*	0.0918	0.033
	(0.113)	(0.113)	(0.115)	(0.115)
Access to extension	0.3502***	0.114***	0.0633	0.023
	(0.118)	(0.118)	(0.113)	(0.113)
Radio ownership	0.0851	0.027	0.1530	0.0554
	(0.156)	(0.156)	(0.148)	(0.148)
Mobile phone ownership	0.3860	0.125	-0.1628	-0.059
	(0.291)	(0.291)	(0.262)	(0.262)
Membership in credit and saving groups	0.0308	0.01	-0.2570**	-0.09**
	(0.113)	(0.113)	(0.115)	(0.115)
Membership in cooperatives	0.4490***	0.145***	0.0224	0.008
	(0.126)	(0.126)	(0.117)	(0.117)
Poundability	-0.0845	-0.027	0.1762***	0.064***
	(0.054)	(0.054)	(0.055)	(0.055)
Quality of garri	0.2144**	0.069**	0.1464**	0.053**
	(0.088)	(0.088)	(0.030)	(0.030)
Quality of fufu	0.1776**	0.057**	0.0257**	0.0009**
	(0.081)	(0.081)	(0.011)	(0.001)
Ease of loading	0.0950*	0.031*	0.0442	0.016
	(0.054)	(0.054)	(0.053)	(0.053)
Starch content	-0.0091	-0.0029	-0.0028	-0.001
	(0.055)	(0.055)	(0.054)	(0.054)
Tapioca taste	-0.1326***	-0.043***	0.0170*	0.006*
	(0.046)	(0.046)	(0.009)	(0.009)
Market demand	0.0139	0.0045	0.0987	0.036
	(0.083)	(0.083)	(0.087)	(0.087)
Root size	0.1650	0.053	0.0547	0.02
	(0.109)	(0.109)	(0.112)	(0.112)
White color	-0.0562	-0.018	-0.0411	-0.015
	(0.057)	(0.057)	(0.086)	(0.086)
Sores well	0.1264*	0.041*	0.0599	0.022
	(0.067)	(0.067)	(0.084)	(0.084)
Root yield	0.0199***	0.0064***	0.4633***	0.168***
	(0.0011)	(0.0011)	(0.122)	(0.122)
Pest and disease resistance	-0.1044	-0.034	-0.1524**	-0.055**
	(0.068)	(0.068)	(0.067)	(0.067)
Stem yield	-0.0122	-0.004	0.2082***	0.075***
	(0.060)	(0.060)	(0.063)	(0.063)
Early maturity	0.1600*	0.052*	0.1796*	0.065*
	(0.087)	(0.087)	(0.095)	(0.095)

Village fixed effects included but not reported. Standard errors clustered at enumeration area level are reported in parenthesis, *** p<0.01, ** p<0.05, * p<0.1

when using responses from the household heads. As such studies that assume a unitary (male) decision-maker decides whether or not to adopt improved cassava varieties as the basis of their analysis may miss some underlying factors. The importance of each factor on the adoption decision of the household head/ spouse depends on the relative importance of the spouse/household head in agricultural-related activities. This is obviously clear as factors such as ease of loading are only significant for household heads (this activity is mainly undertaken by a male). Similarly poundability, an activity usually performed by women, becomes significant only for spouses. By incorporating indicators of household structure and spouses bargaining power, adoption decision at household level can therefore be estimated more precisely.
Like preference heterogeneity, differences in main access variables such as extension, credit, and membership in social networks between the spouse and the household head affect adoption decision differently. However, it must be noted that since the household head is more influential than the spouse (in terms of at least access to basic financial and information services), the effects on the household head may have to be taken as a basis for interpreting the results. For example, access to extension and credit do not have any significant effect on the adoption of improved cassava varieties from the responses of the spouses. However, these variables are significant when the responses from the household head are used in the regression analysis. Using estimated coefficients from the household head, access to credit and extension increases the probability of adoption of improved cassava varieties by 6.7% and 11.4%, respectively.

Reference

Abdoulaye, T., A. Abass, B. Maziya-Dixon, G. Tarawali, R. Okechukwu, J. Rusike, A. Alene, V. Manyong, and B. Ayedun. 2014. Awareness and adoption of improved cassava varieties and processing technologies in Nigeria. Journal of Development and Agricultural Economics 6(2: 67–75.

Aldashev, G., G. Kirchsteiger, and A. Sebald. 2016. Assignment procedure biases in randomized policy experiments. The Economic Journal. DOI: 10.1111/ecoj.12321

Alene, A.D., M. Khonje, T. Abdoulaye, P. Kulakow, and V.M. Manyong. 2012. Adoption of improved cassava varieties in Southwestern Nigeria, IITA Technical Report submitted for the DIIVA project, IITA,Ibadan, Nigeria.

Alene, A.D. 2010. Productivity growth and the effects of R&D in African agriculture. Agricultural Economics 41: 223–238.

Alene, A.D., V.M. Manyong, E. Tollens, and S. Abele. 2007. Targeting agricultural research based on potential impacts on poverty reduction: strategic program priorities by agro-ecological zone in Nigeria. Food Policy 32(3): 394–412.

Alene, A.D. and O. Coulibaly.2009. The impact of agricultural research on productivity and poverty in sub-Saharan Africa. Food Policy 34(2): 198–209.

Bajari, P..and C. Benkard. 2005. Demand estimation with heterogeneous consumers and unobserved product characteristics: a hedonic approach. Journal of Political Economy 113(6):1239–1276.

Bandiera, O. and I. Rasul. 2006. Social networks and technology adoption in northern Mozambique. The Economic Journal 116(514): 869–902.

Barrett, C. 1996. On price risk and the inverse farm size-productivity relationship. Journal of Development Economics 51(2): 193–215.

Barrett, C., F. Bellemare, and J.Y. Hou. 2010. Reconsidering conventional explanations of the inverse productivity size relationship. World Development 38(1): 88–97.

Bulte, E., G. Beekman, S. Di Falco, J. Hella, and P. Lei. 2014. Behavioral responses and the impact of new agricultural technologies: evidence from a double-blind field experiment in Tanzania. American Journal of Agricultural Economics 96(3): 813–830.

Carletto, C., S. Savastano, and A. Zezza. 2013. Fact or artifact: the impact of measurement errors on the farm size productivity relationship. Journal of Development Economics 103: 254–261.

Chassang, S, E. Snowberg, B. Seymour, and C. Bowles. 2015. Accounting for behavior in treatment effects: new applications for blind trials. PLoS ONE 10(6): e0127227. doi:10.1371/journal.pone.0127227

Collier, P. 1983. Malfunctioning of African rural factor markets: theory and a Kenyan example. Oxford Bulletin of Economics and Statistics 45(2): 141–172.

Conley T. and C. Udry.2010. Learning about a new technology: pineapple in Ghana. American Economic Review 100: 35–69.

Croppenstedt, A., M. Demeke, and M. Meschi. 2003. Technology adoption in the presence of constraints: the case of fertilizer demand in Ethiopia. Review of Development Economics 7(1): 58–70.

Davis, K. 2008. Extension in sub-Saharan Africa: overview and assessment of past and current models, and future prospects. Journal of International Agricultural Education and Extension 15 (3):15–28.

De Janvry, A., A. Dustan, and E. Sadoulet. 2011. Recent advances in impact analysis methods for ex-post impact assessments of agricultural technology: options for the CGIAR. Consultative Group for International Agricultural Research, Independent Science and Partnership Council.

Dellaporta, S.L., J. Wood, and J.B. Hicks. 1983. A plant DNA mini-preparation Version II. Plant Molecular Biology Reporter 1:19–21.

Dercon, S..and L. Christiaensen. 2011. Consumption risk, technology adoption and poverty traps: evidence from Ethiopia. Journal of Development Economics 96(2): 159–173.

Dillon, A., S. Gourlay, K. McGee, and G. Oseni. 2016. Land measurements bias and its empirical implications: evidence from a validation exercise in Nigeria. Available online at: http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/EXTL SMS/0,,contentMDK:23506659~pagePK:64168445~piPK:64168309~theSitePK:3358997,00.html

Elshire, R.J., J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, E.S. Buckler, and S.E. Mitchell. 2011. A robust, simple Genotyping-by-Sequencing (GBS) approach for high diversity species. PLoS ONE 6(5): e19379.

Feder G., R. Just, and D. Zilberman.1985. Adoption of agricultural innovations in developing countries: asurvey. Economic Development and Cultural Change 33: 255–298.
Glaubitz, J.C., T.M. Casstevens, F. Lu, J. Harriman, R.J. Elshire, Q. Sun, and E.S. Buckler. 2014. TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. PLoS ONE 9(2): e90346.
Griliches, Z. 1957. Hybrid corn: an exploration in the economics of technological change. Econometrica, 25(4): 501–522.
Hasselt, M. and C.R. Bollinger. 2012. Binary misclassification and identification in regression models. Economics Letters 115 (2012): 81–84.
Hensher, D., J. Louviere, and J. Swait J. 1999. Combining sources of preference data. Journal of Econometrics 89: 197–221.
Holden, S. and M. Fisher. 2013. Can area measurement error explain the inverse farm size productivity relationship? Centre for Land Tenure Studies Working Paper 12/13.
Koundouri, P., C. Nauges, and V. Tzouvelekas. 2006. Technology adoption under production uncertainty: theory and application to irrigation technology. American Journal of Agricultural Economics 88(3): 657–670.
Ladd G., and V. Suvannunt. 1976. A model of consumer goods' characteristics. American Journal of Agricultural Economics 58(3): 504–510.
Lewbel, A. 2007. Estimation average treatment effects with misclassification. Econometrica 75(2): 537–551.
Maertens, A.and C. Barrett. 2013. Measuring social networks' effects on agricultural technology adoption. American Journal of Agricultural Economics 95(2): 353–359.
Manski, C. 2004. Measuring expectations. Econometrica 72(5): 1329–1376.
Maredia, M., and B. Reyes. 2015. Improving the methods of measuring varietal adoption by farmers in developing countries: Recent experience with the use of DNA fingerprinting and implications for tracking adoption and assessing impacts. 29th International Conference of Agricultural Economists, Milan, Italy, 9–14August 2015.
Munshi, K. 2004. Social learning in a heterogeneous population: Technology diffusion in the Indian Green Revolution. Journal of Development Economics 73(1): 185–213.
Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC ., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. American Journal of Human Genetics. 81(3): 559—575.
Rabbi, I, M. Hamblin, M. Gedil, P. Kulakow, M. Ferguson, A.S. Ikpan, D. Ly, and J.-L. Jannink. 2014. Genetic mapping using genotyping-by-sequencing in the clonally propagated cassava. Crop Science 54(4): 1384–1396.
Rogers, E. 1995. Diffusion of innovations (4th ed.). The Free Press, New York:
Suri, T. 2011. Selection and comparative advantage in technology adoption. Econometrica 79(1): 159–209.
Useche, P., B. Barham, and J. Foltz. 2009. Integrating technology traits and producer heterogeneity: amixed-multinomial model of genetically modified corn adoption. American Journal of Agricultural Economics 91(2): 444–461.
Useche, P., B. Barham, and J. Foltz. 2013. Trait-based adoption models using ex-ante and ex-post approaches. American Journal of Agricultural Economics 95(2): 332–338.
Van Zyl, J., H. Binswanger and C. Thirtle. 1995. The relationship between farm size and efficiency in South African agriculture. World Bank Policy Research Working Paper, 1548 . Washington, DC, USA.
Wossen, T., T. Berger, T. Mequaninte, and B. Alamirew. 2013. Social network effects on the adoption of sustainable natural resource management practices in Ethiopia. International Journal of Sustainable Development & World Ecology 20(6): 477–483.
Wossen T, T. Berger , and S. Di Falco. 2015. Social capital, risk preference and adoption of improved farm land management practices in Ethiopia. Agricultural Economics 46: 1–18.
Wossen, T., S. Di Falco, T. Berger, and W. McClain. 2016. You are not alone: social capital and risk exposure in rural Ethiopia. Food Security, doi:10.1007/s12571-016-0587-5
Young, H.P. 2009. Innovation diffusion in heterogeneous populations: contagion, social influence, and social learning. American Economic Review 99(5): 1899–1924.

Appendix

Village level results

This section presents some descriptive statistics for village level dynamics relevant for the CMS project.

Table 39. Land use characteristics.

Variable	Pooled (n=491)	North (n=125)	Southwest (n=125)	Southeast (n=123)	South-South (n=118)
Land use changes of CROP LAND in the village in the last 10 years	%				
Increased	59.3	67.2	51.2	62.6	55.9
Decreased	35.6	28.8	48.8	35.8	28.8
Same	5.1	4.0		1.6	15.3
Land use changes of FOREST LAND in the village in the last 10 years	Pooled	North	Southwest	Southeast	South-South
Increased	13.6	20.8	32.0	.8	
Decreased	74.1	64.0	65.6	90.2	77.1
Same	12.2	15.2	2.4	8.9	22.9
Land use changes of GRAZING LAND in the village in the last 10 years	Pooled	North	Southwest	Southeast	South-South
Increased	13.8	12.8	27.2	6.6	4.3
Decreased	50.6	59.2	43.2	68.0	17.4
Same	35.6	28.0	29.6	25.4	78.3

Table 40. Socioeconomic characteristics of the community.

	Mean
Distance to the nearest cassava stems market in km (n= 236)	18.03 (60.97)
Adult cut-off age(n= 491)	24.98 (12.07)
Distance to the local village market in km(n= 487)	3 (5.58)
Means of transport most frequently get to the local market(n= 438)	3.28 (2.14)
Distance to the nearest main(district) market in km(n= 490)	12.51 (17.67)
Number of months the road to main market is passable for vehicles in a year (n= 488)	10.81 (5.1)
Cost of transporting 100kg of cassava to the nearest main market (Naira) (n= 489)	399.82 (436.99)
Average one-way transport cost(per person) to the main market using a car (naira/person) (n= 490)	167.67 (174.56)
Distance to the nearest seed dealer in km (n= 479)	12.87 (22.73)
Distance to the nearest fertilizer dealer in km (n= 482)	12.24 (18.43)
Distance to the nearest herbicides/pesticides dealer in km (n= 482)	12.02 (17.72)
Distance to the nearest farmer cooperative society office in km (n= 476)	14.85 (49.34)
Distance to the nearest farmer group/club in km (n= 472)	11.43 (19.03)

Table 41. Level of importance of cassava.

Cassava importance	Pooled	North	Southwest	Southeast	South-South
moderately important	0.6	2.4			
important	5.5	10.4	6.5	2.4	2.6
very important	93.9	87.2	93.5	97.6	97.4

Table 42. Participation in cassava project.

Participate in cassava project	Pooled	North	Southwest	Southeast	South-South
yes	16.8	12.8	17.9	16.3	20.5
no	83.2	87.2	82.1	83.7	79.5

Table 43. Assessment of the socioeconomic status of the community.

Socioeconomic status of your community	Pooled	North	Southwest	Southeast	South-South
The poorest	5.8	6.4	3.4	3.3	10.5
Below average	29.7	16.0	46.2	18.7	39.5
Average	52.6	61.6	39.5	63.4	44.7
Above average	10.8	13.6	10.9	14.6	3.5
The richest	1.0	2.4	0.0	0.0	1.8

Table 44. Availability of cassava stems market.

Are cassava stems sold in the village	Pooled	North	Southwest	Southeast	South-South
Yes	45.1	30.3	12.2	84.4	54.5
No	54.9	69.7	87.8	15.6	45.5

Table 45. Means of transport to the local market.

Means of transport most frequently get to the local market	Pooled	North	Southwest	Southeast	South-South
Walking	39.3	16.8	37.8	51.2	47.4
Bicycle	7.8	17.8	5.1	.8	8.6
Vehicle	22.1	34.7	49.0	6.5	5.2
Other specify	30.8	30.7	8.2	41.5	38.8

Table 46. Quality of roads to the main market.

Quality of road to the main market	Pooled	North	Southwest	Southeast	South-South
Very poor	10.7	5.6	13.1	13.8	10.4
Poor	18.6	20.0	24.6	10.6	19.1
Average	25.8	36.8	30.3	15.4	20.0
Good	23.7	24.8	16.4	18.7	35.7
Very good	21.2	12.8	15.6	41.5	14.8

Table 47: Changes in the cassava based system in the last five years.

Cassava production volume	Pooled(n=487)
Increased	70.8
Decreased	25.1
Same	4.1
Cassava processed products	Pooled((n=487)
Increased	45.6
Decreased	9.2
Same	45.2
Number of cassava varieties	Pooled((n=486)
Increased	66.5
Decreased	11.1
Same	22.4
Prices of cassava products	Pooled((n=487)
Increased	43.5
Decreased	41.9
Same	14.6
Number of farmers growing cassava	Pooled((n=486)
Increased	81.1
Decreased	17.7
Same	1.2
Price of cassava roots	Pooled((n=483)
Increased	40.2
Decreased	41.8
Same	18.0
Cost of hiring labor	Pooled((n=487)
Increased	88.1
Decreased	9.4
Same	2.5
Access to tractor	Pooled((n=352)
Increased	6.0
Decreased	41.8
Same	52.3
Cost of hiring tractor/ha	Pooled((n=353)
Increased	32.9
Decreased	16.4
Same	50.7
Cassava yield changes/ha	Pooled((n=482)
Increased	62.9
Decreased	29.7
Same	7.5
Cassava land area/ha	Pooled((n=486)
Increased	70.6
Decreased	25.1
Same	4.3
Use of cassava as source of income	Pooled((n=484)
Increased	69.6
Decreased	23.6
Same	6.8
Prevalence of cassava disease/pests	Pooled((n=486)
Increased	41.6
Decreased	40.9
Same	17.5
Use of external inputs on cassava farms such as inorganic fertilizer and herbicides	Pooled((n=473)
Increased	50.1
Decreased	39.1
Same	10.8

Table 48. Average price of cassava roots and stems.

Dry Season	2015	2014	2013	2012	2011
Cassava root price (₦/metric ton)	10,106.07	14,320.98	13,824.98	12,626.22	11,690.31
Cassava stems price (₦/bundle)	200.3	265.9	225.7	190.3	166.9
Rainy Season					
Cassava root price (₦/metric ton)	10,872.81	11,145.64	11,096.77	10,381.23	95,11.81
Cassava stems price (₦/bundle)	310.5	291.4	268.4	232.2	219.4

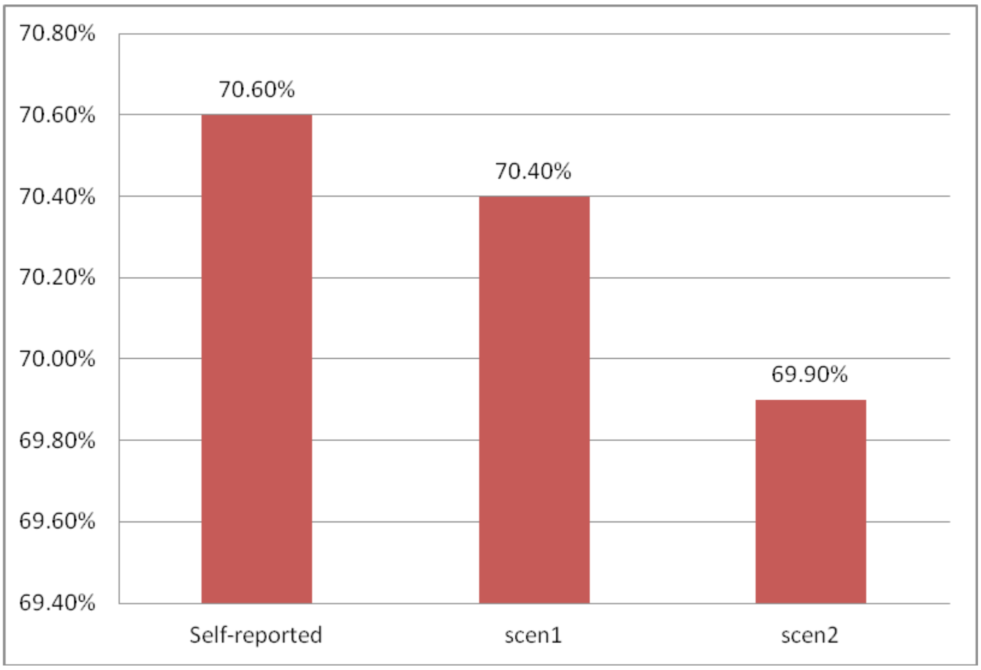


Figure 20. Intensity of adoption among adopters.

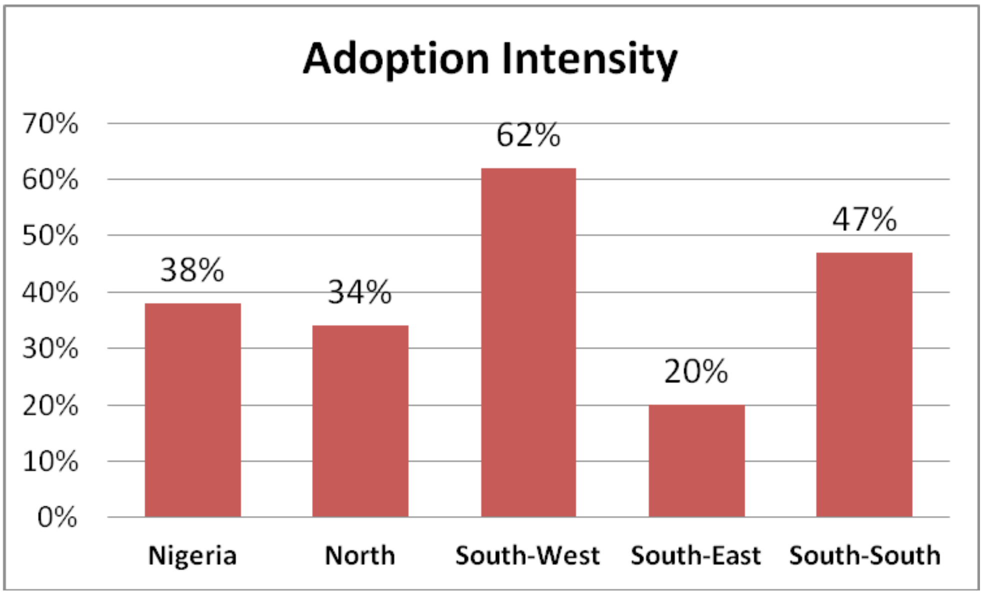


Figure 21. Intensity of adoption among adopters.