Analysis of cassava varietal adoption in Vietnam using DNA fingerprinting approach[1]

Dung Phuong Le[2], Ricardo A. Labarta[3] and Mywish K. Maredia[4]

Abstract

This paper analyzes the adoption level and determinants of different improved cassava varieties in Vietnam. The data come from a nationally representative household survey of 949 cassava families located in 79 villages across the country. We employed DNA fingerprinting approach for better identifying cassava varieties in farmer fields and reducing measurement bias in the regression analysis. Using multivariate probit model with regional fixed effect, the results reveal strong correlation between fertilizer applications, cassava usage and the adoption of improved varieties (IV). Besides, fertilizer intensification, access to credit and selling directly to starch company are important factors influencing farmer's choice to adopt modern improved varieties instead of old improved varieties.

Keywords: Cassava, technology adoption, DNA fingerprinting, multivariate probit, Vietnam

---

[1] **This paper is preliminary draft, please do not cite or distribute**

[2] International Center for Tropical Agriculture (CIAT)
[3] International Center for Tropical Agriculture (CIAT)
[4] Michigan State University

1. Introduction

Cassava (Manihot esculenta Crantz) is an important food source in many regions in the world, especially in tropical countries in Africa and Latin America. Originally coming from South America (Allem, 2002; Olsen and Schaal, 2001), cassava was believed to be planted in Vietnam at the end of 18^{th} Century (Tran, Tran, Hoang, & Kawano, 1996). Nowadays, cassava serves as a cash crop for farmers in the country thanks to its use for the starch and ethanol industry. The average yield of cassava has been increasing rapidly during the past 15 years, from approximately 8 tons per hectare in 2000 to approximately 19 tons per hectare in 2015 (FAOSTAT), as a result of adopting improved agricultural technologies including fertilizer and high-yielding varieties.

Over the year, researchers have been conducting a vast of studies trying to investigate the level and determinants of technology adoption among farmers and measure the impact of the adoption to their livelihood (i.e. (Krishnan & Patnam, 2014, Simtowe et al., 2011). While the fertilizer adoption level is relatively straightforward, the estimation of improved variety adoption has traditionally relied on expert opinion, farmer's elicitation and morphological descriptors (Rabbi et al., 2015). However, these methods maintain uncertainty levels which are creating bias and standard errors of the adoption estimates (Maredia et al., 2016).

A few number of recent studies have been trying to overcome the challenge of varietal identification by employing the DNA fingerprinting analysis in socioeconomic adoption researches. Maredia et al.(2016) used data from two pilot studies for cassava in Ghana and bean in Zambia to test alternative approaches of varietal identification and validation against DNA fingerprinting. The methods tested in this study included a number of different techniques grouped into two types: farmer elicitation and expert elicitation methods. The results showed that those methods provided varied adoption estimates of improved varieties and all of them sustained both type I (i.e., local varieties were incorrectly identified as improved varieties) and type II errors (i.e., improved varieties incorrectly identified as improved varieties). The study suggested to apply DNA fingerprinting as the only credible approach for varietal identification, although the scale depends on the available budget. On the other hand, Floro et al. (2017) used DNA fingerprinting results as a benchmark to identify the household determinants of the adoption of improved cassava varieties in Cauca, Colombia. This study found that farmers overestimated their use of improved cassava varieties which leads to measurement bias when using regression analysis to discover the factors influencing farmers' adoption decision. Again, the paper emphasized the important of accurately identifying improved varieties and reducing the misclassification of the dependent variable.

This paper distinguishes itself by providing the analysis of cassava varietal adoption and determinants using DNA fingerprinting method for a nationally representative data on cassava households. Building upon previous pilot studies, this is one of the first studies integrating DNA fingerprinting approach to track varietal adoption as part of representative household survey. We are able to identify adoption rates of different landrace and improved cassava varieties across Vietnam at national level and then make comparison between farmer-reported and geneticist-identified varietal adoption.

Another motivation for the study stems from the nature of cassava sector in Vietnam where more than 90% of cassava area is covered by improved varieties. This is relatively different from previous cassava adoption studies conducted in Africa and Latin America where landrace varieties are remaining a significant player in the sector. The question here is not about the common case of technology adoption being below expectation, but more about understanding the drivers of farmers' choice in the context where they are opener to new technology. The study focuses on the use and determinants of landrace varieties and different improved variety categories. By conducting a multivariate probit regression, the study investigate the factors affecting farmers' decision in moving from landrace varieties to old varieties and to modern varieties.

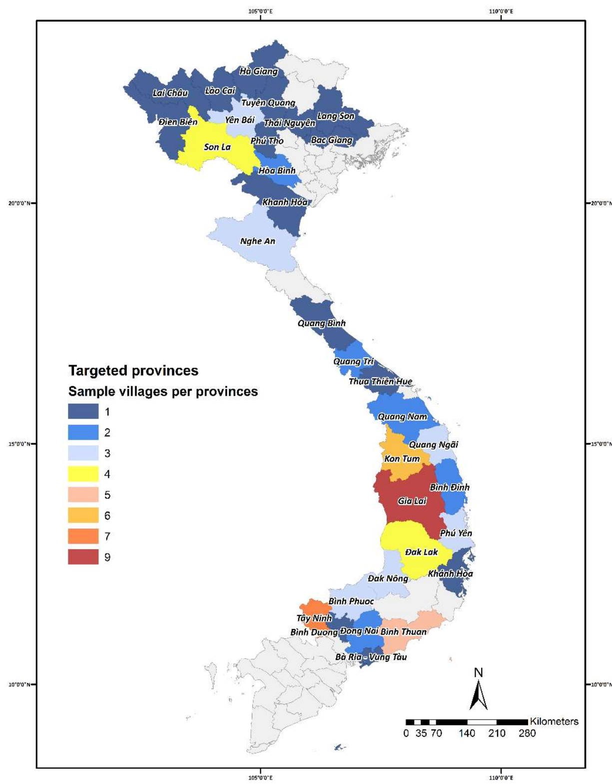
The remaining of the paper is organized as follows. The next section provides an overview of the survey design and data collection. Section 3 presents a brief description of DNA fingerprinting methods and regression model used in our paper. Section 4 reports and discusses the varietal adoption and estimation results. Finally, Section 5 draws conclusion.

2. Survey design and data collection

We collected a national representative sample of cassava household in Vietnam, following a sample design with multistage procedure. First, 32 out of 64 provinces and cities in Vietnam were selected, representing 95% of the cassava area in the country. We conducted power calculation to identify the number of villages and households for sampling, then used probability proportional to size sampling method to calculate the number of village samples in each provinces. The idea was to select randomly 1 villages in each community and randomly interview 12 households in each village with cassava production. However, some villages did not have 12 households producing cassava at the time of the survey, so the team had to visit additional households in neighboring villages. In total, 949 households in 79 villages were involved in the first survey round. In the second round, the same household samples continued to be targeted for follow-up information. Figure 1 presents the distribution of surveyed villages across Vietnam.

After several meetings and trainings in order to design an appropriate questionnaire and pilot surveys from September to October 2015, the official first round of surveys took place between October and December 2015. The survey teams included enumerators from Agricultural Genetics Institute, Thai Nguyen University of Agriculture and Forestry in the North and Institute of Agricultural Science for Southern Vietnam in the South. During this data collection period, besides farmers self-reported cassava varieties and yields, we also collected cassava planting material samples for DNA extraction and fingerprinting. The latter included samples for each farmer-recognized variety (n = 1) for all households and one within-field samples (n= 15) for one or two random fields per village. The second survey round took place from March to May 2016 and aimed to collect follow-up information of the harvested cassava plots. In both phases, different information on household socioeconomic characteristics, agricultural practices, incomes and expenditures, social networks, climate change etc. were also collected. The on-farm data collection are combined and compared with improved methods for varietal identification using molecular fingerprinting.

Figure 1. Sampling distribution across Vietnam



3. Methodological approach

3.1 DNA fingerprinting analysis

The cassava stakes collected in the first round survey were brought to Hanoi for DNA extraction at the Agricultural Genetics Institute, following CTAB-based DNA extraction protocol described by Doyle and Doyle (1990) with minor modifications. Then more than 3300 DNA samples in this study were sent to Cali, Colombia for DNA fingerprinting analysis, including 1570 household samples, 1318 intra-plot samples and 422 institutional collection samples. The institutional samples comes from three key institutions working on cassava varietal breeding in Vietnam: the Root Crops Research and Development Center (RCRDC), the Hung Loc Research and Development Center (HLRDC) and the Agricultural Genetics Institute (AGI). The collection together with CIAT genebank are serving as the reference library to identify the variety of household and intra-plot samples.

The samples were processed using a newly developed protocol for 96 single nucleotide polymorphisms (SNPs) genotyping in cassava with the EP1 system and SNP type assays of

Fluidigm, application version 3.1.2. The technique allowed simultaneously to collect both end-point and real-time data from a unique chip cell with 97% confidence. Quality test included ECU72 genotyping in each chip. From 96 SNP markers belonging to the Cassava Chip-Array, 94 were successful for genotyping. We obtained up to 310.000 data point from 3310 samples. (Floro et al, 2017) provides detail description of the DNA fingerprinting process. Annex A presents a dendrogram of genetic relationship among cassava accessions.

3.2 Conceptual framework and model used

The paper use the M-equation multivariate probit regression econometric model (Greene, 2003) to identify the drivers of different variety adoption in Vietnam:

$$Y_{im}^{*} = \beta_m X_{im} + \epsilon_{im}, m = 1, \ldots., M$$

Where $Y_{im} = 1$ if $Y_{im}^{*} > 0$, 0 otherwise

$Y_{im}(m = 1, \ldots, M)$ represent the unobserved latent variable of cassava varieties adopted by the i^{th} farmer ($i = 1, \ldots, n$). DNA fingerprinting results indicated that there were 85 different varieties planting in farmer′s fields, therefore for simplification and better comparison, we divide the varieties into 3 groups: *landrace varieties, improved varieties released before 2000 (old IV), and improved varieties released after 2000 (modern IV)*. In addition, this classification allows us to understand farmer's decision whether to adopt the old or more recent IV.

X_{im} is a 1xM vector of observed variables that affect the decision of varietal adoption. ϵ_{im}, $m = 1, \ldots, M$ are error terms distributed as multivariate normal, each with a mean of zero, and variance-covariance matrix V, where V has values of 1 on the leading diagonal and correlation $\rho_{jk} = \rho_{kj}$ as off-diagonal elements.

This system of equations is jointly estimated using maximum likelihood method. The model allows us to relax a relatively strict assumption that household's choice of cultivar type to grow does not affect the prediction of the same household's probability of adopting another type. Additionally, we can test if the correlation coefficients among the error terms in each model are significant that will give us the information on the farmer's decision of using more than one variety.

4. Results and discussions

4.1. Socioeconomic characteristics of the sample households

Table 1 presents the means of selected variables by region and adoption categories (1=adopters and 0 otherwise). When breaking down the data into regional characterization, we are able to observe some differences between the five agro- ecological regions: North, North Central, Central Highland, Central Coastal and South East. Particularly, statistics for North, North Central and Central Coastal show that in average, the cassava land area and total land area in these regions are significantly smaller than those in Central Highland and South regions. However, the social network seems to be more active among households in the North, North Central, and Central Coastal than other two regions with 95% household samples have member joining any type of associations. Meanwhile, the Central Highland has lower level of household head education and age than other regions in average. Famers in Central Highland have smaller number of cassava plots but higher number of adopted varieties. Level of fertilizer used in this region is strikingly lower, only 33 percent farmer adopted fertilizer on their cassava plots. The South East region has the biggest land size and the highest number of cassava plots but a bit lower level of seed diversity in average than Central Highland and North Central. The differences in household characteristics between regions give some indications of the geographical effects on cassava production.

Table 1. Descriptive statistics

Variable	All	Region					Adoption category					
		North	North Central	Central Highland	Central Coastal	South	Landrace		Old IV		Modern IV	
		(N = 216)	(N = 132)	(N = 264)	(N = 157)	(N = 180)	Not-planting (N = 690)	Planting (N = 259)	Non-adopters (N = 299)	Adopters (N = 650)	Non-adopters (N = 709)	Adopters (N = 240)
Household size	4.2	4.51	3.85	4.76	3.91	4.34	4.21	4.19	4.66	4.03	4.16	4.47
Gender of household head (1,0)	0.91	0.86	0.95	0.97	0.92	0.93	0.9	0.93	0.92	0.91	0.91	0.95
Education of the HH head	6.78	6.4	7.63	4.75	6.39	7.79	7.12	6.24	6.2	6.99	6.78	6.74
Age of the HH head	50.17	49.32	52.53	42.9	51.77	48.58	51.3	48.36	48.29	50.87	50.35	48.92
Total land size (hectare)	1.62	1.37	0.88	2.76	1.56	5.71	1.73	1.44	2.22	1.39	1.34	3.5
Number of cassava plots	1.8	1.64	1.84	1.43	2.32	1.91	1.76	1.87	1.71	1.84	1.76	2.09
Total cassava land (hectare)	0.72	0.4	0.37	1.49	0.81	3.18	0.86	0.49	0.93	0.64	0.52	2.04
Number of different cassava varieties	1.35	1.26	1.46	1.54	1.17	1.32	1.16	1.66	1.22	1.4	1.32	1.54
Fertilizer application (1,0)	0.86	0.86	0.93	0.33	0.97	1	0.9	0.81	0.82	0.88	0.85	0.95
Agricultural credit (1,0)	0.34	0.35	0.35	0.36	0.24	0.41	0.36	0.31	0.39	0.32	0.33	0.44
Distance to extension office (km)	4.41	4.11	4.34	8.2	3.14	4.39	3.72	5.53	5.57	3.96	4.29	5.17
Social network (1,0)	0.89	0.93	0.97	0.69	0.85	0.58	0.87	0.91	0.84	0.91	0.91	0.72

The DNA fingerprinting analysis found 85 different varieties from the farmers' household samples in the study. Of 949 in total, 259 households remained planting landrace varieties in their land. Some of these also adopted improved variety, others only planted landrace seed. In average, household head of families using landrace varieties are less educated than that of households not using landrace varieties. Households with landrace varieties also have smaller total land and cassava land size. Besides, the descriptive result confirm that old improved varieties releasing before 2000 maintains as favorable seeds among farmers with almost 68 percent adopters, much higher than landrace or modern improved varieties. Adopters of old IV have higher education but lower total land and cassava land size. The trend is clearer in the case of modern improved varieties such that the land size is much bigger and fertilizer application rate is higher in average for adopters in this case. The socioeconomic descriptive statistics are in line with some literature which found bigger land size induces higher probability to adopt new agricultural technology due to better capacity to purchase improved technologies and greater ability to cope with the loss if the technology fails(i.e. (Floro et. al., 2017), (Katengeza et al., 2012)).

4.2. The area adoption rates of cassava varieties in Vietnam

The correct identification of varieties plays a crucial role to study the impact of improved varieties on farm productivity and farm income. In this study, we collected the information on the varietal adoption by interviewing farmer and then confirming with DNA fingerprinting method. The data from farmers' self-identification show that many farmers use their local adapted names to report their cassava variety. We found 120 unique names in total, each unique name was assumed to present a farmer's reported variety. The most common name is "high yielding" cassava ("*cao san*" in Vietnamese) which is widely used for many different varieties and represents a cultivar group. Some varieties were named based on their appearance such as "purple sprout cassava" ("*san dot tim*"), "bamboo leaf cassava" (*san la tre*). Farmers also name their cassava after the original source where they acquired the variety "Tay Ninh high-yield" ("*cao san Tay Ninh*"), "Dong Nai variety" ("*giong Dong Nai*"), "Binh Dinh cassava" ("*san Binh Dinh*"). In some areas, farmers are more familiar with the breeder's name of the cassava. For example, "KM 94" is reported by a majority of respondents in My Hiep commune, Binh Dinh province and Thuong Am commune, Tuyen Quang province. Given the variation in the way farmers name their cassava variety, using farmer's self-reported variety names to estimate varietal adoption can be challenging. It is also not easy to identify clearly which are improved varieties and which are local varieties. Importantly, it is easy to mix so-called cultivars groups with unique varieties.

The High Yielding group has the highest adoption rate with 22.2 percent of total adoption area. The next one is Red Ear which is more popular in the Southern of Vietnam and can be KM419 according to some national breeders. KM94 is the only breeder's variety names appeared in the top six most common variety names provided by farmers. The adoption rate of this variety is 7.8 percent which is much lower than estimated in previous studies (75.14 percent area adoption reported in (Robinson & Srinivasan, 2013). This can be explained by the fact that not many farmers know the breeder name, instead they use local adapted names (i.e High Yielding, Cut). Table 2 presents the area adoption rate of different varieties identified by farmers in comparison with the adoption rate identified by DNA fingerprinting approach. Besides, prior to this study, another study conducted by CIAT tried to document the adoption level of cassava varieties in Vietnam by using expert opinion method. We also included the adoption results from the cassava experts in Table 2 for comparison purpose.

Using cassava stake collected in farmer field for DNA analysis, we were able to identify a more accurate adoption rate of cassava varieties. The reference library allowed us to match each genotype group with the breeder's cultivar sample. Column 3 in Table 2 presents the estimates of the area adoption rate of different cassava varieties identified by DNA fingerprinting analysis. The results show that out of total 85 different varieties identified in farmers' field, KM419 gains the highest adoption rate with 38.73 percent area. KM94 remains as one of the dominant varieties planted in Vietnam with 31.6 percent which is significantly lower than expectation of previous studies and expert opinion. Before, many experts believe that KM94 was the most popular variety planted in Vietnam with 60 percent area adoption rate (Column 2). Landrace varieties with 59 different genotype groups accounted for 9.38 percent cassava area. The results confirmed the domination of improved varieties in Vietnam that was reported in numerous adoption studies (Robinson & Srinivasan, 2013).

Table 2. Estimated area adoption rate of cassava varieties by expert elicitation, farmer survey and DNA fingerprinting methods

Expert elicitation of % cassava area		**Farmer-reported of % cassava area (English -translated names)**		**DNA fingerprinting of % cassava area**	
KM 94	60.00%	High Yielding	22.20%	KM419	38.73%
KM 140	16.30%	Red Ear	13.90%	KM94	31.67%
KM 98-5	4.40%	New	8.60%	KM101	7.98%
KM 419	4.10%	Breeding	7.90%	KM140	3.99%
KM 60	3.26%	KM94	7.80%	KM60	1.30%
Rayong 72	2.76%	Cut	3.50%	Other IV	6.95%
Other improved Varities	3.98%	Other IV	32.60%		
Traditional/Local/Landraces	5.20%	Traditional/Local/Landraces	3.50%	Traditional/Local/Landraces	9.38%

4.3. Determinants of cassava varietal adoption

The multivariate probit system was jointly estimated for three dependent variables: Landrace, Old improved varieties and Modern improved varieties. The random variates drawn is 50 times to ensure the good estimation of the correlation coefficients as suggested by (Cappellari & Jenkins, 2003). The regression results are presented in Table 3. The Wald test provides the P-value of 0.0000 indicating that the model is highly significant. We include the regional fixed effect in the model to control for unobservable factors that would lead to omitted variable bias problem.

The results show that there are a number of factors significantly affecting farmers' decision on cassava varieties when controlling for the regional fixed effect. The household socioeconomic characteristics affect differently on the adoption decision of each variety type. Education of household head increases the adoption of all IV types and decreases the adoption of landrace varieties, though not significantly. Gender of household head only significantly affect the choice of landrace varieties. Having female household head reduces the use of landrace varieties, or in the other word, relatively increases the use of improved varieties. Many recent studies have investigated the relationship between gender and technology adoption and found similar result. For instance, (Lamprecht 2015) suggested that women played a key role in cassava production in Vietnam and they owned better knowledge and sensitivity toward the recognition of different cassava traits. (Seymour, Doss, Marenya, Meinzen-Dick, & Passarelli, 2016) found mixed evidence on whether or not women's empowerment positively associated with higher adoption of improved maize varieties, however, they showed that women's empowerment was positively correlated with greater participation by women in different aspect of improved maize varieties production. Besides, age of household head is also negatively correlated with the choice of landrace varieties.

The variables related to cassava production included in the model are *log cassava land, share of rental land, number of plot, number of varieties, fertilizer, chemical fertilizer per hectare, and access to agricultural credit.* Cassava land size has negatively significant impact on the adoption of old IV, however, the share of rental land positively correlated with old IV adoption??? Number of varieties positively affect the choice of all varietal types with high level of significant at 1 percent because the more varieties the more likely farmers plant different cassava types. The plot number is also positively correlated with landrace varieties and modern IV. As forecasted, the use of fertilizer is negatively associated with the use of landrace varieties but positively associated with the IV adoption. Looking at the intensity of fertilizer application, measured by amount of chemical fertilizer per hectare, we find that farmers using modern varieties tend to associate with more intensive chemical fertilizer application. This is consistent with a number of literature studying the relationship between the adoption of fertilizer and IV (i.e.(Emerick, Janvry, Sadoulet, & Dar, 2016). In addition, our model shows that farmers having access to agricultural credit is also more likely to adopt modern IV at the 10 percent significant level. Fertilizer intensity and access to credits are two of the key factors determining the difference of farmer's choice between old and modern IV.

Table 3. Multivariate probit estimation of the adoption determinants

VARIABLES	Landrace	Old IV	Modern IV
Female-headed	-0.537**	0.193	-0.280
	(0.223)	(0.206)	(0.208)
Age of household head	-0.00865*	0.000904	-0.00164
	(0.00494)	(0.00423)	(0.00522)
Education of household head	-0.0166	0.0157	0.00296
	(0.0186)	(0.0165)	(0.0184)
Household size	-0.00155	-0.0319	0.0378
	(0.0387)	(0.0353)	(0.0369)
Share of dependents	-0.000512	0.000771	-0.00391
	(0.00236)	(0.00224)	(0.00270)
Log of cassava land	-0.00409	-0.108**	-0.0241
	(0.0526)	(0.0484)	(0.0566)
Share of rental land	-0.00614	0.00378*	0.000205
	(0.00392)	(0.00214)	(0.00238)
Number of plots	0.175***	-0.00956	0.244***
	(0.0523)	(0.0534)	(0.0637)
Number of varieties	0.681***	0.500***	0.700***
	(0.113)	(0.111)	(0.114)
Applying fertilizer	-0.674***	0.367**	0.482***
	(0.187)	(0.178)	(0.173)
Fertilizer intensification	-6.56e-05	-9.94e-05**	0.000203*
	(6.06e-05)	(4.86e-05)	(0.000114)
Access to agricultural credit	-0.174	-0.145	0.197*
	(0.115)	(0.103)	(0.120)
Member of organization	-0.0202	0.420***	-0.179
	(0.149)	(0.116)	(0.140)
Off-farm income	0.0518	-0.0822	-0.0317
	(0.114)	(0.103)	(0.118)
Cassava sold direct to starch company	-0.0902	-0.889***	1.415***
	(0.216)	(0.178)	(0.178)
Cassava for selling	-0.685***	1.262***	0.812*
	(0.233)	(0.255)	(0.430)
Cassava for human consumption	1.206***	-0.568***	-1.736***
	(0.195)	(0.178)	(0.409)
Fixed effect: North Central	-0.860***	1.027***	0.0189
	(0.181)	(0.194)	(0.231)
Fixed effect: Central Highland	-1.540***	1.572***	0.384*
	(0.213)	(0.212)	(0.201)
Fixed effect: Central Coastal	-2.061***	1.073***	0.211
	(0.268)	(0.178)	(0.210)
Fixed effect:South	-1.293***	-0.294	1.570***
	(0.245)	(0.181)	(0.243)
Constant	0.745	-1.311***	-3.714***
	(0.496)	(0.476)	(0.685)
Observations	913	913	913

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1
Log pseudolikelihood = -861.72157 Wald chi2(66) = 609.48 Prob > chi2 = 0.0000

In addition, we find that other aspects including institutional membership and cassava usage matter. The result suggest that having household member joining at least one organization increases the use of old IV. Regarding cassava utilization, we investigate the effect of different common purposes including human consumption, livestock feed and commercial. We find strongly positive correlation between human consumption and the use of landrace varieties, yet negative correlation between the same variable and the use of old and modern IV. In contrast, farmers selling their cassava either as fresh root or dried chip are more likely to adopt old and modern IV and less likely to use landrace varieties. This is the evidence that farmers in Vietnam are maintaining the landrace varieties for consumption and using more IV as cash crop. In term of commercialization, we also look at whether or not farmer directly sell their cassava to the starch company and find significantly different results. Farmers selling directly to the starch company are more likely using modern IV and less likely using old IV. In the survey, farmer explained that selling to starch company is more beneficial because they can have better price compared to selling through intermediaries.

Table 4 presents the estimation of correlation coefficients of error terms between three models. The correlation coefficient estimates are jointly significant at 1 percent level indicating the evidence of strong association in farmer's decision of different varietal adoption. The result find negatively significant correlation between old IV and landrace varieties, and between old IV and modern IV. Meanwhile, the correlation between modern IV and landrace is not significant.

Table 4. Correlation coefficient between the equations

Pairs	Correlation	Standard error	t-value	P>|z|
Old IV - Landrace	-0.68869	0.049924	-13.79	0.000
Modern IV - Landrace	0.029772	0.083708	0.36	0.722
Modern IV - Old IV	-0.65841	0.058913	-11.18	0.000

Likelihood ratio test of rho21 = rho31 = rho32 = 0: chi2(3) = 248.567 Prob > chi2 = 0.0000

5. Conclusion

Cassava is serving as a key cash crop for many families living in rural of Vietnam. Majority of cassava varieties growing across the country is improved varieties due to its superior yield and starch content relatively to landrace varieties. In this paper, we have identified the national adoption rate of different cultivars by mobilizing DNA fingerprinting method. We also estimates the determinants of varietal choices among farmers in Vietnam based on a multivariate probit model.

The study finds that farmer's elicitation method failed to identify the adoption level of different cassava varieties in Vietnam. Farmers use their local-adapted names and often mix up between the cultivar groups and unique variety. Using DNA fingerprinting through SNPs for stake samples taking from the farmer fields, we are able to know exactly the variety planted and document the adoption level of each individual variety. We find a relatively large diversity with 85 different cassava cultivars grown in Vietnam. Nevertheless, the results are conditional to the sampling method which assumes that farmers are aware of the number of cassava varieties in their fields. This assumption might need to be tested in future adoption studies.

The regression result indicates the strong correlation between fertilizer applications, cassava usage and the adoption of improved varieties. Besides, fertilizer intensification, access to credit and selling directly to starch company are key variables influencing farmers' decision to adopt modern improved varieties instead of old improved varieties. We also find significant regional effect on the adoption of different cassava varieties. These results suggest the variables one should pay attention when working on the dissemination and adoption of cassava varieties.

References

Allem, A. C. (2002). The origins and taxonomy of cassava. In *Cassava: biology, production and utilization* (pp. 1–16). Wallingford: CABI. http://doi.org/10.1079/9780851995243.0001

Cappellari, L., & Jenkins, S. P. (2003). Multivariate probit regression using simulated maximum likelihood. *The Stata Journal*, 3(3), 278–294.

Emerick, K., Janvry, A. De, Sadoulet, E., & Dar, M. H. (2016). Technological innovations, downside risk, and the modernization of agriculture. *American Economic Review.* 106(1), 1537–1561. http://doi.org/10.1257/aer.20150474

Floro,V. O. IV, R. Labarta, L.A. Becerra, J.M. Martinez, T. Ovalle. (2017) Household determinants of the adoption of improved cassava varieties using DNA fingerprinting to identify varieties in farmer fields: A case study in Colombia. *Journal of Agricultural Economics* forthcoming.

Greene, W. H. (2003). *Econometric Analysis.* (P. Education, Ed.)Journal of the American Statistical Association (Vol. 97). Prentice Hall. Retrieved from http://pubs.amstat.org/doi/abs/10.1198/jasa.2002.s458

Katengeza, S. P., Mangisoni, J. H., Kassie, G. T., Sutcliffe, C., Langyintuo, A., Rovere, R. La, & Mwangi, W. (2012). Drivers of improved maize variety adoption in drought prone areas of Malawi. *Journal of Development and Agricultural Economics*, 4(14), 393–403. http://doi.org/10.5897/JDAE12.029

Krishnan, P., & Patnam, M. (2014). Neighbors and extension agents in ethiopia: Who matters more for technology adoption? *American Journal of Agricultural Economics*, 96(1), 308–327. http://doi.org/10.1093/ajae/aat017

Lamprecht, M., & Lamprecht, M. (2015). Genetic diversity and farmers' selection of cassava (Manihot esculenta Crantz) varieties on small-scale farms in Northern and Central Vietnam.

Maredia, M. K., Reyes, B. A., Manu-Aduening, J., Dankyi, A., Hamazakaza, P., Muimui, K., Raatz, B. (2016). Testing Alternative Methods of Varietal Identification Using Dna Fingerprinting:

Results of Pilot Studies in Ghana and Zambia. *MSU International Development Working Paper*, 149(149), 1–36. Retrieved from http://fsg.afre.msu.edu/papers/idwp149.pdf

Rabbi, I. Y., Kulakow, P. A., Manu-Aduening, J. A., Dankyi, A. A., Asibuo, J. Y., Parkes, E. Y., … Maredia, M. K. (2015). Tracking crop varieties using genotyping-by-sequencing markers: a case study using cassava (Manihot esculenta Crantz). *BMC Genetics*, 16(1), 115. http://doi.org/10.1186/s12863-015-0273-1

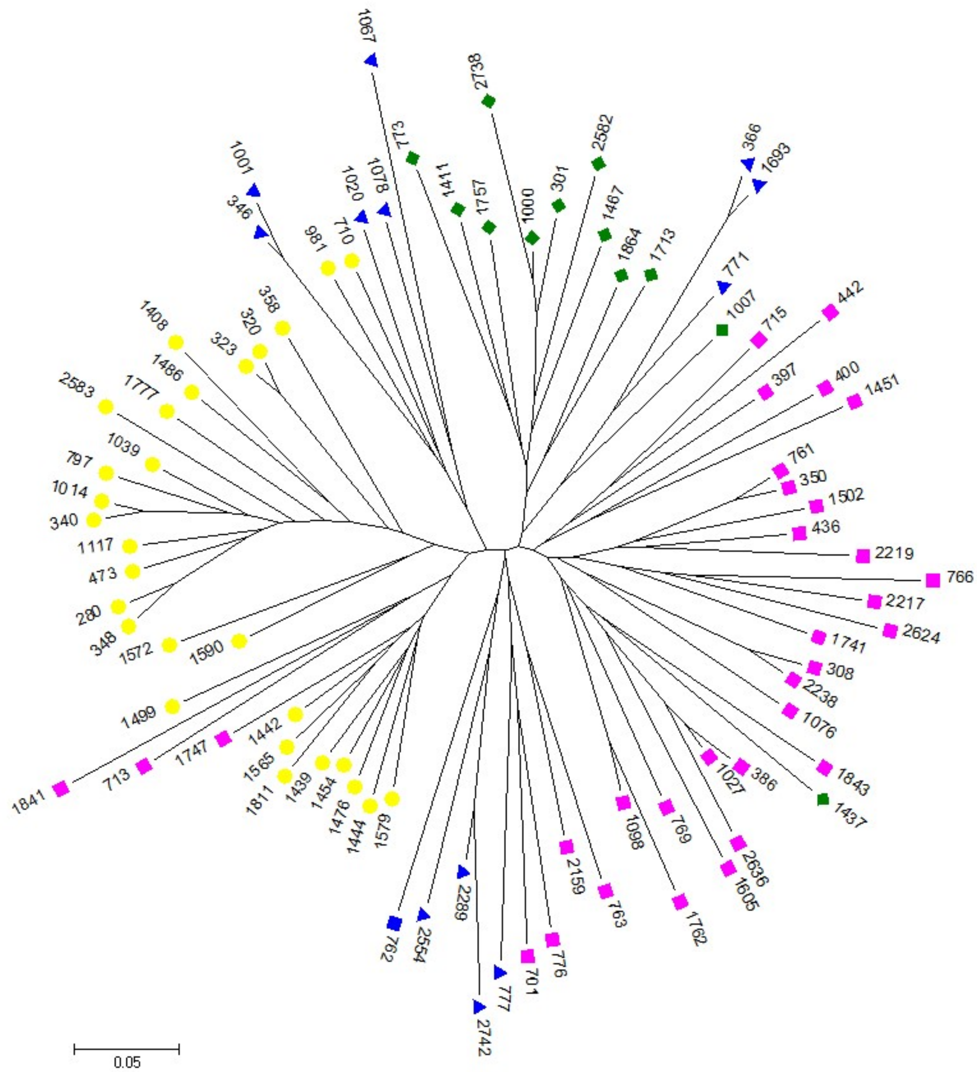
Robinson, J., & Srinivasan, C. (2013). Case-studies on the impact of germplasm collection , conservation , characterization and evaluation (GCCCE) in the CGIAR Case-studies on the impact of germplasm collection , conservation , characterization and evaluation (GCCCE) in the CGIAR Foreword, (September).

Seymour, G., Doss, C., Marenya, P., Meinzen-Dick, R., & Passarelli, S. (2016). Women's Empowerment and the Adoption of Improved Maize Varieties: Evidence from Ethiopia, Kenya, and Tanzania. Agricultural & Applied Economics Association Annual Meeting, 1–30.

Simtowe, F., Kassie, M., Diagne, A., Asfaw, S., Shiferaw, B., Silim, S., & Muange, E. (2011). Determinants of agricultural technology adoption: The case of improved pigeonpea varieties in tanzania. *Quarterly Journal of International Agriculture*, 50(4), 325–345.

Tran, N., Tran, N., Hoang, K., & Kawano, K. (1996). Cassava cultivars and breeding research in Vietnam. Retrieved from https://cgspace.cgiar.org/handle/10568/54823

Annex 1. Dendrogram of identified cassava varieties in Vietnam



Explanatory note: Dendrogram of 85 household samples representing 85 varieties based on the share-allele genetic distance calculated from data of 94 SNP markers, using the neighbor-joining algorithm as the clustering method. Each color represents one subpopulation of cassava in Vietnam

Annex 2. List of improved variety identified in the household samples and their year of first use

Variety	Year of first use	Category
KM101	1989	Old improved varieties
KM94	1993	Old improved varieties
KM60	1993	Old improved varieties
SM 937.26	1994	Old improved varieties
KM98-7	1998	Old improved varieties
11Sa05	Unidentified	Old improved varieties
11Sa08	Unidentified	Old improved varieties
13Sa04	Unidentified	Old improved varieties
HB60	Unidentified	Old improved varieties
KM229	Unidentified	Old improved varieties
SM2	Unidentified	Old improved varieties
SM9326-11	Unidentified	Old improved varieties
Sa21-12	2001	Modern improved varieties
KM419	2006	Modern improved varieties
KM140	2007	Modern improved varieties
Sa06	2008	Modern improved varieties
HL-S11	2010	Modern improved varieties