IIIS Discussion Paper

No.411/ October 2012

The Unbanked Four-Fifths: Informality and Barriers to Financial Services in Nigeria

Michael King
Trinity College Dublin

IIIS Discussion Paper No. 411

The Unbanked Four-Fifths: Informality and Barriers to Financial Services in Nigeria

Michael King[1]
Trinity College Dublin

[1] I would like to acknowledge Darrell Beghin and Andrea van der Westhuizen, both of Finscope and Robert Stone and Jack Willis, both of Oxford Policy Management for their generous assistance with the data. I would also like to thank and Patrick Honohan (Central Bank of Ireland, Carol Newman (Trinity College Dublin), Thorsten Beck (Tilburg University), Ciara Walsh (University College Dublin) and Pedro Vicente (Universidade Nova de Lisboa) for helpful comments.

The Unbanked Four-Fifths: Informality and Barriers to Financial Services in Nigeria

Michael King[1]
Trinity College Dublin

Abstract

Four-fifths of the adult population do not have access to formal financial services in Nigeria. This paper examines the characteristics of 'unbanked' households in Nigeria, investigates the extent and determinants of supply and demand side barriers, and explores the specific role played by informality in financial exclusion. There is evidence to suggest that in Nigeria the unbanked four-fifths have lower incomes, lower education, are less likely to have a mobile phone, and have lower levels of financial sector knowledge and formal documents in their name than the remainder of the population. A non-mutually exclusive framework for analysing barriers to formal banking is developed and insights on the characteristics of sub-groups of the unbanked population who face particular barriers are provided. While poverty and distance to bank branch are the most important barriers cited, informality also plays a significant role and tackling informality directly represents an opportunity for financial inclusion policy. Using instrumental variables, the precise role played by informality in financial exclusion is estimated and for individuals with four and five documents, it is found that an additional document increases the probability of being banked by 17 percent and 15 percent respectively.

[1] I would like to acknowledge Darrell Beghin and Andrea van der Westhuizen, both of Finscope and Robert Stone and Jack Willis, both of Oxford Policy Management for their generous assistance with the data. I would also like to thank and Patrick Honohan (Central Bank of Ireland, Carol Newman (Trinity College Dublin), Thorsten Beck (Tilburg University), Ciara Walsh (University College Dublin) and Pedro Vicente (Universidade Nova de Lisboa) for helpful comments.

1. Introduction

Financial exclusion remains a global phenomenon with up to 2.5 billion people globally excluded from the formal banking system (Morduch et al, 2009). In Africa, estimates from the Finscope surveys suggest that the proportion of the population without access to formal financial services ranges from 44 percent in South Africa to 92 percent in Mozambique, with Nigeria, the most populous country in Africa, lying at the higher end of this scale with 79 percent, approximately four-fifths of the adult population, estimated to be 'unbanked' (Honohan and King, 2012).[2]

Improving access to financial services is increasingly considered a key objective in wider efforts to reduce poverty and vulnerability in low income countries. Economic theory and an increasing body of empirical research suggest that access to savings, payment and credit services facilitates consumption smoothing, helps insure against risk and allows investment in education and other forms of capital. Despite this, little is known about the relative importance of different barriers to financial services. Disentangling the roles played by demand constraints, such as income insufficiency, poor education, informality and financial illiteracy, and supply constraints, such as distance and high cost is a crucial first step in attempts to design effective policies to broaden the reach of formal financial services.

Recent research estimating access to formal banking services in Mexico, transition economies and Sub-Saharan Africa has shed light on the characteristics of financially excluded households (Djankov, Miranda, Seira and Sharma (2008), Beck and Brown (2010), Honohan and King (2012) and King (2012)). In their study of Mexican households, Djankov, Miranda, Seira and Sharma (2008) provide evidence that increased financial inclusion is important for

[2] Figure 1 in Honohan and King (2012) illustrates the relationship between mean income levels and the percent of the population with access to formal financial services for 10 countries surveyed by Finscope. Nigeria is an outlier, recording a low level of financial inclusion for its relatively high level of mean income among its Sub-Saharan African peers.

economic development when they find that households with bank accounts enjoy higher levels of consumption, possess greater assets and are more likely to be college educated. Similarly, Beck and Brown (2010) find that the likelihood of holding a bank account or bank card increases with income, wealth and education in most transition countries. Honohan and King (2012), in a study on eleven Sub-Saharan African countries, find that while an individual's income, education, and proximity to services help determine the likelihood of having access to financial services, individual level psychometric perspectives also play a significant role. The first contribution of this paper is to extend the approach taken in Honohan and King (2012) by exploiting a richer Nigerian dataset and using a series of state level variables to derive more precise estimates of the determinants of personal financial access.

Recent literature helps provide the analytical context for the treatment of barriers to financial services used in this paper. Beck and de la Torre (2006) provide an analytical framework for understanding barriers to payment, savings and credit services that differentiates between demand and supply insufficiency constraints and a host of underlying conditions or state variables.[3] Claessens (2006) makes a further distinction between voluntary exclusion, defined as having access to but not making use of financial services because of a lack of awareness, price/poverty reasons, or a simple lack of need for the services, and involuntary exclusion, defined as having no access. This paper develops a framework for understanding the barriers faced by individuals in availing of formal financial services. The framework, a series of non-mutually exclusive self-reported supply and demand barriers derived from respondent answers to the Finscope survey, provides evidence of the relative importance of the different barriers from the perspective of the individual respondent, and allows for the assessment of

[3] These include market size, macroeconomic fundamentals, technology, the level and distribution of income and the costs of doing business.

individual and state level determinants for the citing of each barrier. This is the second contribution of this paper.

One such barrier to financial inclusion is individual level informality, a constraint cited by 15 percent of the unbanked population in Nigeria.[4] Eligibility criteria to open a bank account can play a key role in financial exclusion. If an individual fails to meet the documentary requirements to open a bank account increases in income, a fall in the cost of banking or a change in preferences will not allow for inclusion in the banking sector. Informality induced involuntary exclusion can occur for individuals who may demand and have sufficient resources to avail of financial services but who are prevented from access. In this context, greater formalisation could therefore offer the potential to make improvements to financial inclusion, without the need to wait for a wider development process and higher incomes.

While informality is likely to be related to low incomes, low education and distance from administrative centres, it remains a distinct form of poverty. Individuals in the informal sector may enjoy reasonable levels of income, send their children to school and yet they are not particularly formalised. This may be due to recent migration, weak institutions where either the state fails to deliver opportunities to formalise in the form of land rights, voter registration cards etc., or a preference on behalf of the individual to avoid labyrinthine or dysfunctional institutions.

Recent literature has challenged many of traditional perceptions on informality as pre-capitalistic or pre-development enterprises or households. One strand postulates that perverse incentives can encourage transition to informality by previously formal micro-enterprises, particularly in overly-bureaucratic institutional settings (Jonasson, 2011 and Hanson, 2010). A more benign view points to the fact that "survival type" informal enterprises can play a

[4] There is also reason to believe that informality plays a role for the 64 percent of individuals who cite poverty as a reason for financial exclusion.

significant role in reducing poverty (Gulyani and Talukdar, 2010). The determinants of informal micro-enterprises are considered by Loayza, Servén and Sugawara (2009) and they suggest that a combination of poor public services, a burdensome regulatory regime, and weak monitoring and enforcement capacity by the state lead to higher levels of informality.

The theory and empirics behind individual level informality has been explored in the literature in less detail. Bureaucratic barriers to individual level formalisation, coupled with heterogeneous perceptions about potential benefits and mistrust of state institutions means that many of Africa's poor may not actively pursue formalisation. However, formalisation can help reduce poverty and vulnerability through improvements in land, employment and political rights and access to utilities and banking services.

The cross country evidence provided by Beck, Demirguc-Kunt and Martinez Peria (2006) suggests that documentary requirements are a significant barrier to financial services in Nigeria.[5] Beck, Demirguc-Kunt and Martinez Peria (2006) estimate that the number of documents needed to open a checking account in Nigeria is high, estimated at 3.66 official documents, the ninth highest of the 54 developing countries surveyed. However, particular administrative barriers to the formalisation of the urban and rural poor exist in most of sub-Saharan Africa, but are particularly important in Nigeria. According to Davis et al (2005) passports are issued only for approved travel purposes in Nigeria, while driving licenses are

[5] Beck, Demirguc-Kunt and Martinez Peria (2006) figures for Nigeria are averages from a survey completed by three of Nigeria's largest banks. The authors find that the cost of banking services is high in Nigeria. For transactions, although the annual fee for a checking account is low in Nigeria, the minimum balance needed to open a checking account is higher than annual GDP per capita, the second highest of the 55 countries surveyed. In addition, the cost of using an ATM card is particularly high at 60 cents per $100 USD withdrawn. For savings accounts it is a similar story. The minimum balance required to open a savings account is over 20 percent of GDP per capita, the seventh highest of the 55 countries surveyed, while the annual fee for a savings account is negligible. The database from Beck, Demirguc-Kunt and Martinez Peria (2006) suggests that Nigeria has a comparatively high ranking for 'locations to open a deposit account', suggesting that physical access is less of a barrier to access when compared with other developing countries.

available only to drivers, and proof of a registered residential address eludes the population living in rural areas and in informal settlements.

The importance of informality in determining financial access is suggested by the Finscope Nigeria dataset. The 16,127 unbanked respondents have an average of 0.94 official documents in their name, compared with 2.95 for the 4,131 'banked' respondents. Following this motivation, the third and final objective of this paper is to investigate the role played by individual level informality in financial exclusion. Specifically, the hypothesis that informality plays a distinct and significant role in financial exclusion will be tested with an instrumental variables approach.

To overcome the potential endogeneity of informality (measured by the number of formal documents held) and personal access to financial services, two instruments are employed; the type of refuse collection used and the time to water source faced by individual respondents. Controlling for heterogeneity at the local state and local government area level, evidence is found that the number of official documents plays a significant role in the likelihood of being banked. The result is particularly strong for individuals around the threshold of four to five formal documents who can be considered on the cusp of financial inclusion.

The remainder of this paper proceeds as follows: Section 2 describes the data, both the 2008 Finscope Nigeria survey and ancillary data sources. Section 3 presents the methodologies employed and outlines the identification strategy. Section 4 presents the empirical results and Section 5 concludes.

2. Data Description

3.1 Primary Variables of Interest

Individual level data is taken from the 2008 Finscope survey of Nigeria which asks 21,110 individuals across all 36 states plus the Federal Capital Territory (FCT) of Abuja about their use of an array of financial products. The survey also records details about a respondent's personal characteristics, quality of life and psychographics. The sample design and weighting procedure was conducted by the National Bureau of Statistics (NBS) and a random selection of eligible members in each household is conducted by the Kish grid Table Method. The Nigerian dataset is one of four Finscope surveys that oversampled less populated states by surveying approximately 560 individuals in each of the 36 states and the FCT of Abuja. For nationally representative summary statistics and regressions, adult weights are employed and in this paper these weights are used unless otherwise stated.

Table 1 provides the summary statistics and coding for all the individual level variables used in this paper. While binary variables such as *urban*, *mobile phone* and *female* are straightforward, further definition is required for other variables used such as *formally banked, financial status, education*, *personal monthly income, financial sector knowledge*, *bank trust*, *risk aversion*, *formality, time to water source*, *type of refuse collection* and *time to store*.

In the Finscope surveys, respondents are asked whether they currently use or have up to 30 different financial products. While it is obvious for the vast majority of the financial products which of these are formal or informal, there may be a difference of opinion for some of the marginal cases.[6] Two approaches are taken to defining personal financial access in this paper. Following the approach taken in Honohan and King (2012), formally banked is firstly defined

[6] Following the definition of formal bank products used in Honohan and King (2012), Table 1 lists the products and services that define a respondent as being formally or informally banked in this paper.

as those with personal access to formal financial products from formal financial institutions. A slightly modified variable *financial status* is also employed on occasion, defined as the three mutually exclusive statuses: formally banked, informally banked only and financially excluded. Summary statistics for these three groups are presented in Table 2.

Four-fifths of Nigerians surveyed do not have access to formal banking services, with only 3 percent of those surveyed declaring access to formal credit services such as an Islamic loan, an overdraft, a loan from a bank, a mortgage or housing loan, a credit card or a valu card. For those who have access to formal banking services (20.1 percent), the majority have personal access to savings and payment services (19.4 percent), with only 3.1 percent enjoying personal access to formal credit facilities. The descriptive statistics find that people living in a rural location, women and those with low education and income are less likely to have access to formal financial services. Informal access to financial services is more widespread, with 25.2 percent of respondents reporting personal access. Those deemed as financially excluded - individuals with no access to formal or informal savings, transaction or credit services - total 53.8 percent of the population.

The education variable is standardised on a scale of 1-8 from "no formal education" up to "completed university education". As shown in Table 1, the average level of education among respondents lies between primary school completed and some secondary education.

Unlike many of the other Finscope surveys, personal income in Nigeria is recorded as the specific monthly value. I use the average 2008 exchange rate, the year of the survey, to convert these values into US Dollars. The average personal monthly income is $69, and the distribution is skewed heavily to the right.[7]

[7] A significant number of income observations are recorded as zero, which meant that when the natural log of income was used, about 35 percent of the observations were lost. To overcome this issue, $1was added to each respondent's monthly income.

The financial sector knowledge score is a normalised score (on a scale of 1-10) achieved in a financial sector knowledge quiz given during the interview. A series of financial products/terms are mentioned and the individual receives 2 points for “I understand", 1 point for "I have heard of" and 0 points for "never heard of/ don’t understand". With 33 products, there the highest score possible is 66 but as the score is reduced to a one to ten scale, the average respondent has financial sector knowledge of 3.6 out of 10.

To support the analysis I derive two important psychometric variables for each individual; risk aversion and trust in banks. The measure of risk aversion is calculated using an individual’s response to the question: "*To get ahead in life, one needs to take some risks"*. Sixteen percent of respondents are deemed to be risk averse. The banktrust variable is determined by the respondent’s reaction to the statement: *“I trust banks”*. Thirty six percent of respondents confess to trusting banks.

The level of informality among respondents is measured using the question: *“Which of the following documents, if any, do you have that are in your name?”* The list of options include document such as an electricity bill, a water bill, a municipal rates and taxes invoice, a telephone bill/account, a lease or rental agreement, a tax return, an insurance policy document, a national identity card, an international passport, a payslip, land ownership documents or house ownership documents. For each positive response the respondent achieves one point, up to a maximum of 16 points. The average Nigerian enjoys 1.33 formal documents; not sufficient to open a bank account. Figure 1 displays the distribution of the number of official documents for the banked, informally banked only and financially excluded populations.

3.2 Data on Barriers to Formal Banking Services

As noted Claessens (2006) differeniates between voluntary and involuntary exclusion, both demand side issues, and supply side considerations. Building on this approach, a comprehensive framework to classify respondents' self declared reasons for being unbanked is developed. The framework emanates from the survey question *"There are many reasons why people often don't have a bank account. You said earlier that you don't currently have a bank account. Why is this?"* Potential answers range from supply side issues such as 'snobbish staff/not helpful', profit/interest is low to demand side issues such as lack of income or preference to deal in cash. Table 3 presents an overview of the way in which all 32 possible responses are classified in this study. The framework, depicted in Figure 3, illustrates the eleven categories that responses are grouped into. Demand side issues are divided into involuntary and voluntary exclusion. Reasons for being involuntarily excluded include poverty, informality, financial illiteracy and low education, whereas reasons for voluntary exclusion include simple preference or religious values. Supply side barriers include high cost, distance, poor service, lack of trust and safety concerns.

The important point is that the reasons for not having a bank account are not mutually exclusive. A respondent can point to multiple reasons for being unbanked across demand and supply constraints. In total there are approximately 16,000 people without a bank account and over 41,000 constraints to financial access were identified.

The majority of barriers cited by respondents are classified as involuntary exclusion reasons (82 percent of unbanked respondents cite an involuntary constraint), with poverty reasons (64 percent of unbanked respondents), financial illiteracy (24 percent of unbanked respondents) and informality reasons (15 percent of unbanked respondents) representing the largest components. Supply constraints make up the second most important set of reasons cited (36

of unbanked respondents), with high cost (14 percent of unbanked respondents) and distance (10 percent of unbanked respondents) the two most frequently cited reasons for being unbanked. Only 8 percent of unbanked respondents cite reasons pertaining to voluntary exclusion.

3.3 Ancillary Data Sources

The Finscope dataset is complemented with a series of state level financial sector and institutions variables. First, the Oxford Policy Management (OPM) Supply Side Survey 2010 comprises supply side data from a range of financial service providers such as commercial banks, microfinance banks and payment service providers as well as the Central Bank of Nigeria. Most relevant for this paper is the data from six of the largest retail banks in Nigeria. In fact, 63 percent of all those 'banked' in the 2008 Finscope survey held accounts with one of these six banks. I have taken three variables from this dataset: the number of branches per million inhabitants, the number of ATM locations per million inhabitants and the number of microfinance branches per million inhabitants in each state. The microfinance data covers 823 branches of the total 901 regulated by the Central Bank of Nigeria in June 2010. Unsurprisingly, the geographical distribution of bank branches and ATMs is positively related as shown in Figure 2.

The second source of state level data is the 2008 version of Afrobarometer Nigeria. This survey provides state level measures of crime, trust in other Nigerians, the level of corruption (bribery) and formality (pay income taxes) in Nigeria. The Afrobarometer surveyed a nationally representative, random, stratified probability sample of 2,324 Nigerians. For this paper, the average (mean) responses on the four state level variables of interest are employed. First, the degree of state level bribery is derived from the individual question: *"Do you have to give a bribe, give a gift, or do a favor for government officials to get a document or a*

permit, to get water or sanitation service, or to avoid a problem with the police?". A dummy variable *used bribe* is then calculated and the proportion of people who used a bribe in each state is deduced. Second, the question *"How much do you trust each of the following types of people: Other Nigerians?"* is used to derive the dummy variable *societal trust.* Third, whether or not the individual has paid income taxes is used as a proxy for state level formality. Finally, the state level of crime is derived from the proportion of individuals who have experienced crime, directly or within their family, or feared for their safety in their own home over the past year. Table 1 provides the details and descriptive statistics for these state-level variables.

3. Econometric Model and Identification Strategy

3.1 Determinants of Financial Status

The first part of the analysis is to ascertain the determinants of financial status at the individual level, accounting for relevant state variables. In doing so, I build on and modify the model used in Honohan and King (2012). Specifically, I use three probit models to assess the probability of individual i in state n being banked, informally banked only or unbanked (the opposite of banked) as follows:

$$\Pr\{FINANCIAL\ STATUS_{i,n}\} = f(IND_{i,n}, GEO_{i,n}, STATE_n^f, STATE_n^{nf}) \quad (1)$$

where $IND_{i,n}$ are characteristics of the individual respondent; $GEO_{i,n}$ are characteristics of that individual's local environment; $STATE_n^f$ are state level financial services infrastructure variables defined as branches per million, ATMs per million and micro-finance branches per

million; and $STATE_n^{nf}$ are other state characteristics such crime levels, levels of bribery, level of societal trust, state level informality and income per capita.[8]

3.2 Barriers to Formal Financial Services

To understand the characteristics of respondents who face different barriers to formal financial services I employ a multivariate probit model. The approach involves a simulated maximum likelihood (SML) estimator that compares all those who cite a certain constraint, as defined by the various branches of Figure 3, against all other unbanked respondents. While similar to the generic probit specification, the multivariate probit model specifically allows for the estimation several correlated binary outcomes jointly. The model is estimated at two levels; at the level of voluntary exclusion, involuntary exclusion and supply constrained (the upper branches of Figure 3) and then on three separate occasions for each of the disaggregate reasons given for being unbanked under the 'buckets' voluntary exclusion, involuntary exclusion and supply constrained. In each scenario the model is estimated for the unbanked four-fifths of individuals in the dataset.

The multivariate probit regressions are conducted using a maximum likelihood estimator with a Geweke-Hajivassiliou-Keane (GHK) smooth recursive conditioning simulator which enjoys a number of desirable properties such as simulated probabilities that are unbiased, bounded within the (0,1) interval and more efficient in terms of the variance of the estimators of probabilities than other simulators (Borsch-Supan and Hajivassiliou, 1993). The multivariate probit model conducted for the upper branches of Figure 3 can be depicted as follows:

$$\Pr\{CONSTRAINT\ FACED_{c,i,n}\} = f(IND_{i,n}, GEO_{i,n}, STATE_n^f, STATE_n^{nf}) \qquad (2)$$

where c = supply reasons, involuntary exclusion or voluntary exclusion.

[8] Given the mutually exclusive nature of the categories banked, informally banked only and financially excluded, I first used a multinomial logit model but the assumption of the independence of irrelevant alternatives (IIA) was violated.

The analysis is further deepened by using three separate multivariate probit models for the specific constraints cited as one moves down the barriers framework outlined in Section 3 and graphically presented in Figure 3. The multivariate probit approach was chosen because the non-mutually exclusive nature of the responses leads to correlation in the error terms and requires a flexible underlying covariance structure.

3.3 Formal Documents and Financial Status

In the final step of the analysis, I attempt to estimate the precise role played by a respondent's informality, as measured by the number of official documents they possess in their name, in their likelihood of being banked. The 16,127 unbanked respondents in the survey have an average of 0.94 official documents in their name. In comparison, the 4,131 respondents that have bank accounts possess an average of 2.95 official documents in their name.

The initial model I employ is a probit model with the number of documents as one of the explanatory variables. Equation 3 outlines the basic probit approach taken.

$$\Pr\{BANKED_{i,n}\} = f(IND_{i,n}, GEO_{i,n}, STATE_n^f, STATE_n^{nf}) \quad (3)$$

There is reason to suspect that formality, as measured by the number of documents in possession, may not be exogenous, which could lead to misleading estimates. Unobserved heterogeneity at the individual level such as conscientiousness or motivation may make the individual more likely to simultaneously have a high number of official documents and be banked. For example, it is possible that individual i has an innate conscientiousness or administrative abilities, which could make the collection of official documents easier and increase the probability that they engage with formal financial services. There may also be unobserved heterogeneity at the village level in the form of an efficient local administrator or

uncontrolled for strategic proximity considerations that may also simultaneously lead to individual i having a high number of official documents and availing of formal financial services. Either of these types of endogeneity may lead to an over estimation of the role played by level of formality in determining the likelihood of being banked in the probit estimates.

Reverse causality may also be a feature of this model as it may plausibly be argued that registering for a bank account may be an early step in getting other official documents such as utility bills. While the minimum balance to open a checking account in Nigeria, estimated to be 106 percent of GDP per capita for a checking account by Beck, Demirguc-Kunt and Martinez Peria (2008), is likely to reduce the likelihood that individuals would register for a bank account in order to access additional documents and increase their formality, it remains plausible that having a bank account opens up opportunities to formalise. The bias in the probit estimates due to potential reverse causality could run in either direction.

As a result of these two potential biases, an instrument that directly affects the number of official documents individual i possesses but does not directly affect the likelihood of being formally banked, except through the number of documents channel, is required. To achieve this, two instruments, *type of refuse collection* and *time to water source,* are employed.

First, survey respondents are asked what type of refuse collection they use. In increasing degrees of informality, responses include: household bin collected by government or household bin collected by private company (both coded as one); disposal within compound (coded as two) and unauthorised refuse heap or throw in drainage/flowing river (coded as three). The rationale behind using type of refuse collection as an instrument is that it provides us with information regarding the informality of individual i, which I argue should not

independently affect the individuals likelihood of being banked but will affect the number of formal documents the individual possesses.

With significant populations of urban poor as well as rural poor, Nigeria is characterised by significant heterogeneity in housing conditions and access to public utilities (such as refuse collection). It is argued here that individual level informality, as distinct from economic or educational poverty, can be proxied in part by the degree of formality of the type of refuse collection practices engaged in by the individual. An informal method of refuse collection indicates that an individual is less likely to enjoy housing facilities associated with utility bill documents and less likely to own their own home or have a formal lease, leading to a lower number of formal documents.

I find that this instrument enjoys a relevance to the potentially endogenous variable, number of official documents, as evidenced by the significant coefficient (at the 5 percent level) in Table 8, the first stage regression. The relevance of this instrument is tested separately and the first stage of the exactly identified instrumental variables estimation with refuse collection as the only instrument produces an Angrist-Pischke multivariate F-test of excluded instruments of 7.13.

Second, respondents are asked how long it takes to get to their main water source. The variable is coded in increasing time from dwelling as follows: in dwelling (coded as one); less than 5 minutes (coded as two); up to 15 minutes (coded as three); up to 30 minutes (coded as four); up to 1 hour (coded as five) and more than 1 hour (coded as six). The rationale behind using time to water source as an instrument is that it provides us with information regarding the informality of individual i, which I argue should not affect the likelihood of being banked with the inclusion of certain controls.

In similar fashion, time to water source is likely to tell us something about individual level informality. If an individual faces a significant distance to water, they are unlikely to enjoy housing facilities that generate utility bill documents and less likely to own their own home or have a formal lease, leading to a lower number of formal documents.

I find that this instrument displays a high degree of relevance to the potentially endogenous variable, number of official documents, as shown by the significant coefficient in Table 8, the first stage regression. In similar fashion, the relevance of this instrument is tested separately and the first stage of the exactly identified instrumental variables estimation with refuse collection as the only instrument produces an Angrist-Pischke multivariate F-test of excluded instruments of 26.91.

However, in this paper the two instrumental variables are employed in the identification strategy. The first stage results of the instrumental variable estimation with both refuse collection and time to water source as instruments allow us to reject the null hypothesis of weak identification; the Angrist-Pischke multivariate F-test of excluded instruments is 17.12. Possible overidentification is tested for with the Sargan-Hansen test. In this test, the residuals from the 2SLS are regressed on all exogenous variables and the test statistic indicates whether or not it is possible to accept the null hypothesis that all instruments are uncorrelated with the error term. In this model the Hansen J p-value is 0.930, greater than 0.05 and hence the null hypothesis cannot be rejected and it can be concluded that the over identifying restriction is valid.

The IV approach is based on the exclusion restriction that time to water source and type of refuse collection only influence the likelihood of being banked through informality and not through proximity to banking infrastructure. In other words, there is no covariance between

the instruments and the error term in the second stage regression ($Cov\{\varepsilon, Z\} = 0$), where ε is the error term and Z represents the instruments.

Arguably, any weakness in the exclusion restriction is likely to be most relevant in a rural setting where time to water source and formality of refuse collection may be related to proximity to services such as banking infrastructure. However, there may also be other regional characteristics that affect the type of refuse collection or the extent of drinking water infrastructure that simultaneously affect the likelihood of being banked, such as prevailing levels of poverty, informality or the strength of local government.

To overcome these concerns at the state level a series of control variables for average state levels of income, formality, corruption and banking infrastructure are included, as well as a series of state level dummy variables under an alternative specification, and I feel confident that this approach helps surmount the threats to the exclusion restriction at the state level.

Heterogeneity at the village, electoral or local government area level is a more challenging issue to overcome and two approaches are pursued in this regard. First, I include urban/rural dummies and distance to store to at least partially account for the possible relationship between the instruments and distance to banking services. I also endeavour to control for other issues such as prevailing levels of poverty, informality or the strength of local government at the local government level with the inclusion of dummy variables for the 587 local government areas in Nigeria.

A second important argument is the existence of sufficient exogenous variation in our instruments for formality. It was previously argued that informality can be considered a distinct form of poverty which may be related to income, education, location etc, but that it is a distinct form of disadvantage. A number of cross tabulations were conducted to substantiate this claim. First, while the number of documents an individual has is found to have a

consistent relationship with education level, it increases for those in the top three income deciles. On the other hand, time to water source and type for refuse collection do not exhibit variation with education or income decile with an exception for those who have completed university education, a small number of Nigerians, and those in the top income decile.[9] This suggests that there is reason to believe that the two instruments are at least in part reflecting an exogenous component of informality. In any case our specification includes a comprehensive list of control variables including education and income.

Second, it is worth assessing whether either of the instruments varies consistently with time to grocery store, the measure of proximity to services employed in this paper. In cross tabulations, there is no evidence that type of refuse collection varies with time to grocery store but there is modest evidence to suggest that time to water source increases with distance to grocery store. Nevertheless, distance to grocery store is likely to be more related to distance to bank branch and as I include this as a control variable in all specifications, concern over correlation between time to water source and the error term are eased. To reduce the overall violation of the assumption that the instruments are related to the error term I also include a series of interaction terms as further controls.

4. Empirical Results

4.1 Determinants of Financial Status

This paper uses a probit model to estimate the determinants of a respondent's financial status and the results are presented in Table 4. The table reports marginal effects, in other words the change in the probability of being formally banked, informally banked only or unbanked

[9] These cross tabulations are not reported in this paper but are available on request.

for an infinitesimal change in each independent, continuous variable or discrete change for binary or other discrete variables.

The results confirm many of the intuitive conclusions found in Honohan and King (2012).[10] The unbanked four-fifths (shown in the final two columns of Table 4) have lower incomes, lower education, are less likely to have a mobile phone, have lower levels of financial sector knowledge and a lower number of formal documents in their name than the remainder of the population. In terms of psychometric characteristics, the unbanked four-fifths display low levels of trust in banks and higher risk aversion than the population. When state level dummy variables are included, the unbanked group are more likely to be located in rural areas.[11]

In line with the results of Honohan and King (2012), gender is significantly related to measures of financial access only in simpler specifications.[12] The coefficient on female becomes insignificant when the full specification is estimated, in particular when education level and psychometric variables are included.

Respondents who use only informal financial services share a number of characteristics with the unbanked group, such as that they tend to mistrust banks, have low education and have fewer formal documents in their name than the remainder of the population.[13] In terms of monthly personal income, the informally banked group do not suffer from the same level of poverty as the unbanked population. Even when state level controls are included, the informally banked population are associated with states with higher levels of informality, lower numbers of bank branches per million population and higher numbers of atms per million population. This suggests that informal financial services are more likely to be

[10] In Table 4 the comparison group for each group is the complement of that group. For example, the unbanked group is compared with the banked and informal only groups combined.
[11] However, when state level financial sector, institutions and income levels are instead included the this result is insignificant.
[12] These simpler specifications are not represented but are available on request.
[13] The comparison group here is the banked and unbanked groups.

prevalent in states with lower levels of formal banking infrastructure (specifically bank branches), and also higher levels of state level informality that would preclude successful bank expansion strategies based on the traditional documentary requirements to open a bank account.

The probit results for the banked population are reported in columns 1-2 in Table 4. The formally banked population are more likely to live in urban areas, have a high level of bank trust, enjoy higher levels of financial sector knowledge and be less risk averse. They are also likely to have higher income and education, and a greater number of formal documents in their name. Significantly, the banked population live in states characterised by a higher degree of formality, as measured by the proportion of people who pay income taxes and this is likely to be related the presence of more widespread banking infrastructure in states with higher levels of formality.

4.2 Barriers to Formal Financial Services

To assess the determinants of the different barriers to financial inclusion, multivariate probit estimation is conducted for both the high level constraints, voluntary, involuntary and supply barriers, as well as the more specific barriers, represented by the lower branches of Figure 3. It is worth recalling from the methodology section that in each case the analysis is conducted for only the unbanked four-fifths of the sample.

In the analysis, cross contamination is allowed because of the non-mutually exclusive nature of reasons cited for being unbanked, and as a result I find only modest differences between groups, reflecting the multidimensionality of poverty and financial exclusion. Among the unbanked population, unsurprisingly, lower income and lower education increase the likelihood that an individual will cite involuntary exclusion reasons such as poverty, financial

illiteracy; low education and informality (see Table 5). There is evidence to suggest that individuals who cite supply constraints have higher income and a higher number of formal documents, suggesting that to be aware of supply constraints such as high cost, safety concerns and poor service, an individual is likely not to be among the poorest in Nigeria. As distance is the fourth supply constraint, the findings in Table 5 confirm that individuals who cite supply constraints are more likely to be rural dwellers. In addition, those who cite voluntary exclusion reasons tend to have more formal documents but lower financial sector knowledge.

The pooled nature of these categories means that most of the information occurs at the more disaggregated level. As a next step I therefore look at the disaggregated categories of voluntary exclusion, involuntary exclusion and supply constraints in more detail, setting each of the component categories, as represented by the lower branches of Figure 3, as the dependent variable in separate probit regressions and assessing the effect of our various independent variables on the probability of facing these particular constraints. The results, which are shown in Table 6, suggest that individuals who simply prefer, for non-religious reasons, not to engage with formal financial service providers are likely to be more formalised but have lower financial knowledge, when compared with the wider unbanked population. Individuals who cite religious reasons for being unbanked tend to have higher income levels, reflecting the relative economic status of religious groups in Nigeria.

Columns 1-8 in Table 6 present the results for the four voluntary exclusion reasons for being unbanked. While those who cite informality reasons (columns 1 and 2) tend to be educated and rurally based, they do not display any greater levels of informality over the wider unbanked population. This suggests that education helps individuals understand the role played by formality in access to banking services. Poverty is cited as a constraint by 64 percent of the unbanked population and, as shown in columns 3 and 4 of Table 6, these

individuals have lower income and higher financial sector knowledge than the wider unbanked population.

Individuals who cite low education as a reason for being unbanked are characterised by lower levels of education, lower financial knowledge and are less likely to face distances to services. Again confirming the robustness of the results, individuals who cite financial illiteracy as one of their reasons for being unbanked score lower on the financial sector quiz on the questionnaire.

The differences become more obvious when the group of individuals who cite supply reasons for being unbanked is considered as a result of the greater heterogeneity in responses that make up the supply constrained group (see Table 7). Those who cite safety concerns are characterised by lower levels of financial sector knowledge. Those who cite low trust in banks are likely to be more risk averse. Unsurprisingly, distance is cited as a reason for being unbanked more often in rural areas, among those who face longer distances to services.

4.3 Formal Documents and Financial Status

Table 9 presents the results of the probit and Instrumental Variable (IV) estimations of the importance of the number of formal documents in the likelihood of being banked. The probit estimates confirm the prior that the number of formal documents plays a statistically significant role in the determination of being banked. However, as outlined in section 3.3, it could be the case either that the number of formal documents held by an individual is endogenous to being banked or that reverse causality may be present between the number of documents and *formally banked*, or both. With this in mind, columns 3 - 4 in Table 9 provide the IV estimates for the case where the number of formal documents held is instrumented with time to water source and type of refuse collection.

The results show that the number of official documents plays a significant role in the likelihood of being banked. The IV estimates are higher than the probit estimates across all model specifications. The coefficients are 0.54 for each of the two specifications presented. Reverse causality is the most likely reason for this downward bias in the probit estimates.

To help interpret the results, I calculate marginal effects at the mean and at each level of formality for both the probit and IV probit models with state level variables and controls included; the results are presented in Table 10. The marginal effects represent the change in the probability of being banked for a one unit change in the number of formal documents held by the respondent. Similar to the point estimates, the marginal effect at the mean (MEM) is higher after instrumentation, suggesting that a one unit increase in the number of formal documents held by the respondent leads to a 7.7 percent increase in the probability of being banked, although this coefficient is marginally outside the 95 percent confidence level (column 4). This compares with a 2.1 percent estimate from the probit model (column 3).

However, the marginal effect at the mean is not likely to be particularly informative as the change in the probability of being banked is likely to be very different when a respondent is going from two formal documents to three than from nine to ten, or indeed zero to one. Table 10 shows the estimated marginal effects from both models. The change in the distribution in marginal effects between the probit model and the IV probit model is striking. The instrumented marginal effects are higher than the probit estimates between zero and eight formal documents and lower thereafter. Arguably, the instrumented marginal effects are in fact more intuitive as the larger gains from an additional document in terms of the probability of being banked are likely to the higher close to some threshold number of documents, or tipping point, estimated to be at four and five documents. Equally so, increases in the probability of being banked as a respondent goes from eleven formal documents to twelve should be negligible and this is what is found by the instrumented marginal effect estimates.

The analysis specifically finds that individuals who increase their number of formal documents by one when at zero documents increase their probability of being banked by 5.5 percent. However, when an individual possesses four and five documents an additional document increases their probability of being banked by 17 percent and 15 percent respectively (column 2). The result at five documents is statistically significant at the 1 percent level for both specifications with state level variables and state controls.[14]

In other words, the intuitive finding from the marginal effects analysis is that when individuals are close to having sufficient documents, an additional document will lead to a statistically significant increase in the likelihood the individual is banked. However, an explanation for the result that moving from none to one document increases the chances of being banked by 5 percent requires some explanation. It is likely that this result is being driven by the large number of people at this level of documents and it is noteworthy that the size of the coefficient is one third that of the coefficient at the threshold. Indeed, the coefficients for one, two and three documents are each in turn larger than the previous one, suggesting an increasing function towards the threshold, although in these cases the coefficients are not statistically significant.

4.4 Robustness

To assess the robustness of this result, a number of approaches were undertaken. First, a comprehensive array of interaction terms were included in the instrumental variable regressions, specifically all combinations of monthly income, financial sector knowledge, urban, gender, education, banktrust and risk aversion were employed, and the size and significance of the coefficient on number of documents remains remarkably consistent.

[14] As a robustness check I use different versions of the number of formal documents, specifically binary variables for greater than two, greater than three, greater than four, greater than five and greater than six formal documents and find statistically significant coefficients for these variables of interest in line with marginal effects analysis.

Second, I conducted the IV regression separately for both the urban and rural subsamples and find similarly sized coefficients that are again larger than the OLS estimates. This robustness check considered the possibility that our two instruments might be less successful in a rural setting but no evidence was found for such a weakness.

Third, I included local government area dummies in the instrumental variables estimation to control for unobserved heterogeneity at the local level. With a local government area fixed effect linear model I not only find that the two instruments retain their relevance to number of formal documents, and fail to reject the null hypothesis that all instruments are uncorrelated with the error term, but that the coefficient on number of documents remains positive and statistically significant at the 1 percent level.[15] Hence, by taking into consideration local government fixed effects, this suggests that the central result that informality plays a statistically significant and causal role in the determination of banking status is robust.

Fourth, the two categorical instruments employed were transformed into a system of dummy variables and these were used to instrument for formal documents. While none of the individual dummy variables were found to be significant in the first state, which is not surprising, the coefficients on the instrumental variables regression are very similar to the coefficients in the main specification presented. The two coefficients are 0.52 and 0.51 for the specification with state controls and state variables respectively, and both are statistically significant at the 1 percent level. While the marginal effects of this approach are not presented in this paper, they follow almost an identical pattern to the marginal effects presented in Table 10, specifically a threshold effect around four to five documents and an effect moving from no documents to one document. The main difference is that the mean

[15] A weakness of this approach is that the local government areas (LGAs) uniquely identify the dependent variable in a significant number of cases which is not surprising as on average there are only 34 observations for every LGA. A linear model is used as convergence is not achieved for the instrumental variable probit model. As a result a comparable estimate is not generated for inclusion in Table 9.

effect is actually statistically significant at the 1 percent level, whether using state controls or state level variables.

Finally, to assess whether the number of documents was in fact endogenous in the model I use a bootstrapped approach to calculating the Hausman test statistic because of the adherence to robust standard errors. The test results indicate that there is reason to believe that the number of documents held is indeed endogenous to *formally banked*. I reject the null hypothesis of exogeneity, with a p-value of 0.00 when I instrument with *time to water source* and *type of refuse collection*. This suggests that the IV approach was an essential improvement in the estimation.

5. Conclusion

Documentary requirements for opening a bank account are particularly high in Nigeria, estimated at 3.66 official documents (Beck, Demirguc-Kunt and Martinez Peria, 2006). Given that only 9.7 percent of the Nigerian population posses three or more official documents and only 5.6 percent of the population posses four or more official documents, it is unsurprising that four-fifths of the Nigerian population remain outside the formal banking sector.

The evidence on the determinants of the financial status of the Nigerian population suggests that not only are the banked population likely to have more official documents to their name, but they are also more likely to live in states characterised by a higher degree of formality as measured by the proportion of people who pay income taxes.

This paper represents an attempt to understand the relative importance of, and underlying characteristics of the individuals who face different types of barriers to formal banking

services in Nigeria. The results show that significant heterogeneity exists between groups who face different barriers. Policy aimed at improving financial inclusion should adopt differing strategies for excluded populations based on the specific constraints they face. As the data used in this paper emanate from three static surveys with broad based objectives, an opportunity remains open to investigate the nature of barriers to formal financial services through more focused experimental surveys.

The instrumental variable results show that the number of official documents an individual holds plays a significant role in the likelihood of being banked. This result is of particular note as it reveals the possibility that significant increases in financial inclusion may be possible without the need to wait for higher incomes, higher levels of education or greater financial literacy. The result is particularly strong for individuals around the threshold of five to six formal documents who can be considered on the cusp of financial inclusion. Hence, the greater formalisation in the status of the poor in Nigeria, and across Africa is likely to lead to greater use of formal financial services.

6. References

Beck, Thorsten and Martin Brown. 2010. "Which Households use Banks? Evidence from the Transition Economies". European Banking Center Discussion Paper No. 2010–25.

Beck, Thorsten and Asli Demirguç-Kunt. 2008. "Access to Finance: An Unfinished Agenda." *World Bank Economic Review 22(3): 383–396.*

Beck, Thorsten, Asli Demirguç-Kunt and Ross Levine. 2007. "Finance, Inequality and Poverty: Cross Country Evidence." *Journal of Economic Growth,* 12 (1), 211-252.

Beck, Thorsten, Asli Demirguç-Kunt and Maria Soledad Martinez Peria. 2007. "Reaching Out : Access to and Use of Banking Services across Countries." *Journal of Financial Economics,* 85 (1), 234-266.

Beck, Thorsten and Augusto de la Torre. 2006. "The Basic Analytics of Access to Financial Services". World Bank Policy Research Working Paper 4026.

Beck, Thorsten, Asli Demirguc-Kunt and Maria Soledad Martinez Peria. 2006. "Banking Services for Everyone? Barriers to Bank Access and Use around the World". *World Bank Policy Research Working Paper 4079.*

Beck, Thorsten, Asli Demirguç-Kunt and Patrick Honohan. 2009. "Access to Financial Services: Measurement, Impact and Policies". *World Bank Research Observer,* 2009 24(1): 119-145.

Borsch-Supan, Axel and Vassilis Hajivassiliou. 1993. "Smooth unbiased multivariate probability simulators for maximum likelihood estimation of limited dependent variable models," *Journal of Econometrics*, vol. 58(3), pages 347-368.

Claessens, Stijn. 2006. “Access to Financial Services: A Review of the Issues and Public Policy Objectives”. *The World Bank Research Observer 2006 21(2):207-240*.

Djankov, Simeon, Miranda, Pedro, Seira Enrique, and Siddharth Sharma. 2008. “Who Are the Unbanked?” World Bank Policy Research Working Paper 4647.

Galor, Oded and Joseph Zeira, 1993. “Income Distribution and Macroeconomics”. *Review of Economic Studies* 60 (1), 35-52.

Gulyani, Sumila and Debarata Talukdar, 2010 “Inside Informality: The Links Between Poverty, Microenterprises and Living Conditions in Nairobi’s Slums” *World Development* Vol. 38, No. 12, pp. 1710–1726.

Hajivassiliou, Vassilis, Daniel McFadden and Paul Rudd. 1996. "Simulation of multivariate normal rectangle probabilities and their derivatives theoretical and computational results," *Journal of Econometrics*, vol. 72(1-2), pages 85-134.

Hanson, Gordon H. 2010. “Why Isn’t Mexico Rich?” *Journal of Economic Literature*, 48:4, 987–1004.

Honohan, Patrick and Michael King. 2012. “Cause and Effect of Financial Access: Cross-country Evidence from the Finscope Surveys” in Robert Cull, Asli Demirguc-Kunt and Jonathan Morduch (eds.), *Banking the World: Empirical Foundations of Financial Inclusion*. MIT Press, Cambridge.

Honohan, Patrick. 2008a. “Cross-Country Variation in Household Access to Financial Services.” *Journal of Banking and Finance*, 32(11): 2493-2500.

Honohan, Patrick. 2008b. "Household Financial Assets in the Process of Development." in James B. Davies (Ed.), *Personal Wealth from a Global Perspective*. (Oxford University Press).

Honohan, Patrick and Thorsten Beck. 2007. *Making Finance Work for Africa*, Washington, DC: The World Bank.

Honohan, Patrick. 2005. "Measuring Microfinance Access: Building on Existing Cross-country Data." World Bank Policy Research Working Paper No. 3606.

Honohan, Patrick. 2004. "Financial Development, Growth and Poverty: How Close Are the Links?" in Charles Goodhart (Ed.), *Financial Development and Economic Growth: Explaining the Links*. Palgrave: London.

Jonasson, Erik. 2011. "Informal Employment and the Role of Regional Governance" *Review of Development Economics*, 15(3), 429–441.

King, Michael. 2012. "Is Mobile Banking the Tyranny of Distance to Bank Infrastructure? Evidence from Kenya". IIIS Working Paper 412.

Levine, Ross. 2005. "Finance and Growth: Theory and Evidence", in Philippe Aghion and Steven Durlauf (ed.), *Handbook of Economic Growth*, edition 1, volume 1, chapter 12.

Loayza, Norman, Luis Servén and Naotaka Sugawara (2009) "Informality in Latin America and the Caribbean". World Bank Policy Research Working Paper 4888.

Murdoch, Jonathon, and Barbara Haley. 2002. "Analysis of the Effects of Microfinance on Poverty Reduction". NYU Wagner Working Paper No. 1014.

Morduch, Jonathan, Alberto Chaia, Aparna Dalal, Tony Goland and Maria Jose Gonzalez. 2009. "Half the World is Unbanked". Financial Access Initiative Framing Note, October 2009.

Oxford Policy Management (OPM). 2010. "Financial Services Landscape in Nigeria". Summary is available at http://www.efina.org.ng/Financial_services_landscape_summary_15_September.pdf.

Porteous, David. 2007. "Financial Service Access and Usage in Southern and East Africa: What Do Finscope Surveys Tell Us?" August 2007. www.finscope.co.za/documents/2007/CrossCountryreport.pdf

World Bank. 2008. *Finance for All: Policies and Pitfalls in Expanding Access*, Washington, DC: The World Bank.

7. Appendix

Table 1: Summary Statistics and Data Sources

	Coding	Obs	Mean	Std. Dev.	Min	Max	Source
Individual Level Variables							
Financial Status (often the binary variable Formal Banked)	Formally Banked = 1, Informal Only = 2, Financially Excluded = 3. Formally Banked products include An ATM card, Debit card, Credit card, Savings/transaction account, Current or cheque account, Fixed deposit bank account, Mortgage or housing loan, Personal loan from a bank, Vehicle finance, overdraft, Islamic loan and Islamic financing investment. Informal Products include Membership of a saving club, Loan from government, Loan from employer, Loan from a microfinance organisation, Loan from a pool, Loan from a cooperative, Loan from Government to start/run business and Credit from a local store.	21110	2.42	0.8	1	3	FinScope 2008
Formality Documents	1 point given for having each of the 16 forms in respondents name.	21110	1.46	1.89	0	16	FinScope 2008
LnPerIncome	Exact amount turned into US dollars using exchange rate 0.00854 .	21110	69.24	189.96	0	8540	FinScope 2008
Educ	No formal education = 1, Some primary school = 2, Primary school, completed = 3, Some secondary school = 4, Secondary school completed = 5, Professional Qualification or equivalent = 6, Some university = 7, University completed = 8.	21110	3.61	2.14	1	8	FinScope 2008
Age	Exact Age	21110	36.14	14.54	18	99	FinScope 2008
gender	Female = 1, Male = 0	21110	0.48	0.5	0	1	FinScope 2008
Mobile	Yes = 1, No or Not Answered = 0	21110	0.42	0.49	0	1	FinScope 2008
Time to Store	Length of time taken to get to grocery store	21110	15.79	16.61	0	245	FinScope 2008
Urban	Urban = 1, Rural = 0	21110	0.24	0.42	0	1	FinScope 2008
banktrust	Yes = 1, No = 0	21110	0.34	0.47	0	1	FinScope 2008
Risk Aversion	Agree or Don't Know = 0, Disagree = 1	21110	0.16	0.36	0	1	FinScope 2008
FSKNow	Scaled into a 0-10 index. Score given for the following responds for each of the 33 financial terms below: Heard but do not understand = 1 point, Heard and do understand = 2 points.	21109	3.59	2.78	0	10	FinScope 2008
Time to water source (IV)	Coded as follows: in dwelling = 1, less than 5 minutes = 2, up to 15 minutes = 3, up to 30 minutes = 4, up to 1 hour = 5 and more than 1 hour = 6.	21110	2.34	0.89	1	5	FinScope 2008
Type of Refuse Collection (IV)	1 = Household bin collected by government or Household bin collected by private agency, 2 = Disposal within compound, 3 = Unauthorised refuse heap or Throw in drainage/flowing river	21110	2.45	0.65	1	3	FinScope 2008

State Level Variables							
Used bribe	if had to give a bribe or a gift to, or do a favour for government officials to get a document or a permit, to get water or sanitation service, or to avoid a problem with the police: Once or twice, few times, often = 1 (used bribe), otherwise = 0 (did not use bribe)	37	0.32	0.17	0.04	0.81	Afrobaromoter 2008
Societal trust	*How much do you trust each of the following types of people?: Other Nigerians* Dummy variable equal to 1 if answers "I trust them somewhat" or "I trust them a lot. If answers "not at all" or "just a little", set variable equal to 0.	37	0.30	0.21	0.00	0.68	Afrobaromoter 2008
Informality (no income taxes)	The proportion of individuals in the state that pay income taxes.	37	0.75	0.22	0.18	0.98	Afrobaromoter 2008
Crime level	*Over the past year, how often, if ever, have you or anyone in your family: Had something stolen from your house? Been physically attacked? Feared for your safety at home?* Dummy variable equal to 1 if crime or fear of crime experienced and equal to 0 if did not experience or fear crime	37	0.55	0.19	0.14	1.00	Afrobaromoter 2008
Branches per million	Survey of 6 of the largest banks in Nigeria on branch numbers per state	37	12.15	10.53	2.10	55.60	OPM 2010
Atms per million	Survey of 6 of the largest banks in Nigeria on branch numbers per state	37	11.23	11.68	1.80	60.50	OPM 2010
Mfos per million	Survey of 823 Microfinance Organisations in Nigeria	37	5.91	6.10	0.43	29.89	OPM 2010
GDP per Capita	Gross Domestic Product per Capita (in USD and PPP)	37	2124	1289	667	5798	Nigerian Statistics Office

Table 2: Summary Statistics by Financial Status

	Formally Banked	Informally Banked	Financially Excluded
Total	21.1	25.2	53.8
Region			
Urban	39.1	20	40.9
Rural	14.1	27.2	58.7
Gender			
Female	15	27.3	57.8
Male	26.7	23.3	50
Education			
More than completed secondary school	43.7	19.7	36.6
Did not complete primary	2.2	28.1	70
Income[c]			
Under $1 per day	13.2	23.7	63.1
Above $1 per day	31.8	28.2	41.1
Age			
60 +	20.9	19.7	56.2
50-59	22.3	22.9	49
40-49	25.7	28.7	46.3
30-49	23.3	28.1	48.1
16-29	17.9	22.1	60

Note: The unbanked four-fifths are made up of those financially excluded and informal only.

Figure 1: Distribution of Formal Documents

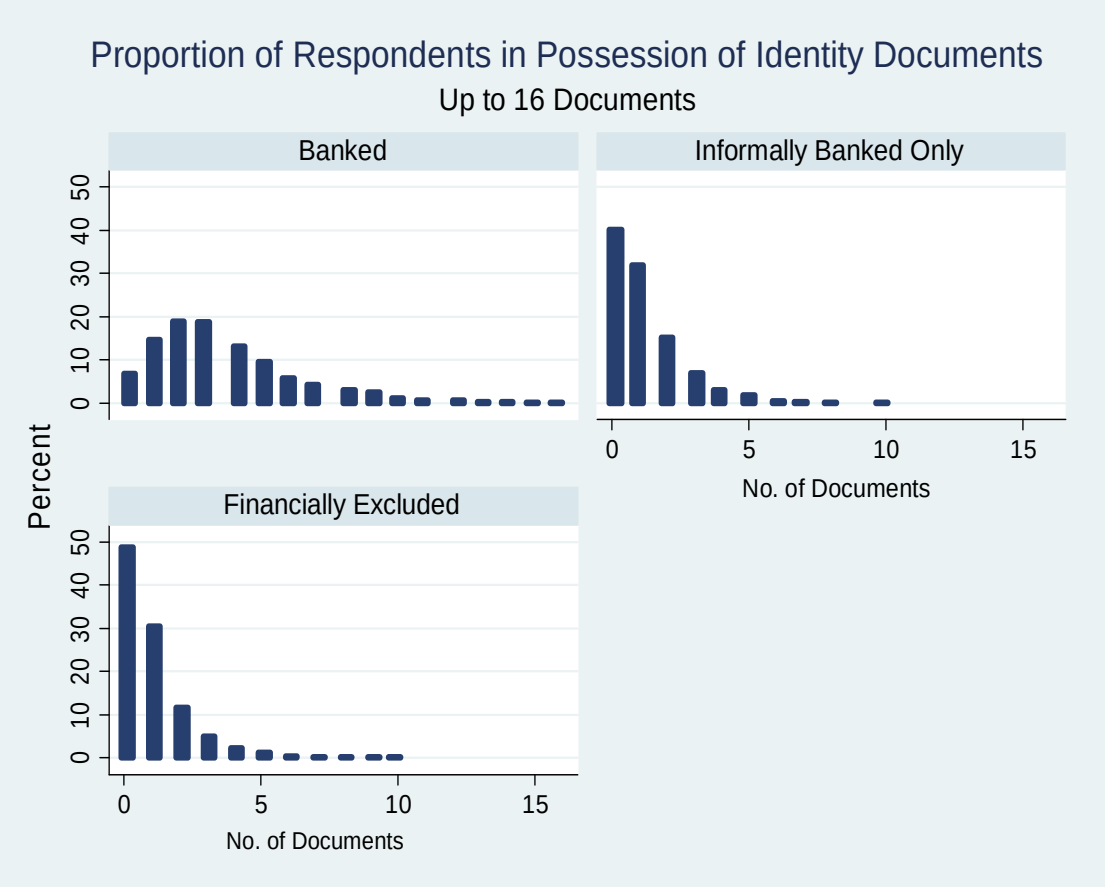

Figure 2: Bank Development by State

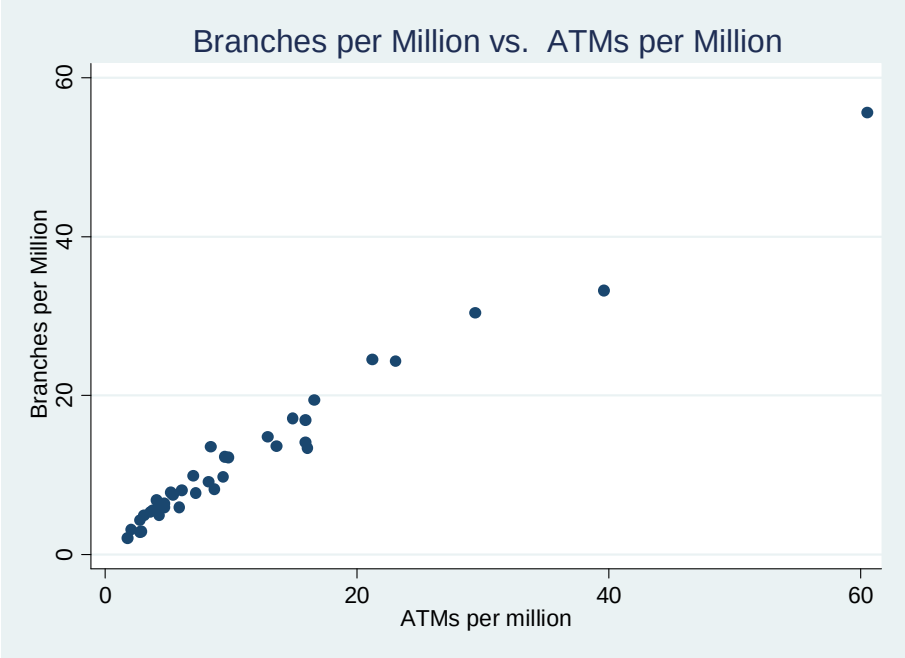

Figure 3: Breakdown of Reasons for Unbanked

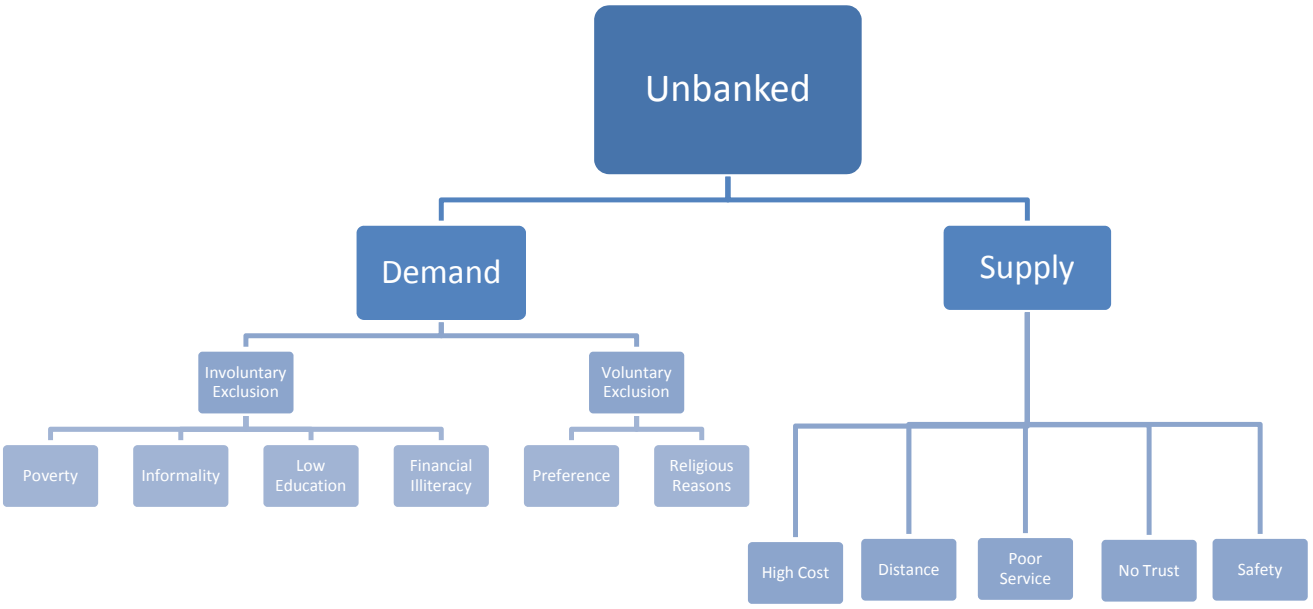

Table 3: Reasons Provided for Not Having a Bank Account, Respondents, Percent[a]					
Barrier (Level 1)	Barrier (Level 2)	Details	Responses	% of Total Responses	% of Unbanked Population
Involuntary Exclusion (Demand Side)	Informality	I don't have an identity document	1,095	2.58%	6.45%
		I don't have a reference for the bank	533	1.26%	3.14%
		Too much documentation involved	1,039	2.45%	6.12%
		Irregular signature	349	0.82%	2.06%
	Poverty	I don't have enough money to open an account	8,318	19.63%	49.00%
		I don't have regular income /Not presently working	4,915	11.60%	28.95%
		Do not have enough collateral	580	1.37%	3.42%
	Financial Illiteracy	Lack of information about bank products and services	2,382	5.62%	14.03%
		Don't know how to open an account	1,898	4.48%	11.18%
	Low Education	Literacy/Can't read or write	2,114	4.99%	12.45%
Demand Voluntary Exclusion	Religious Exclusion	Payment and receiving of interest stops me from going to a bank	280	0.66%	1.65%
	Preference	Never thought about it/ No need for it	1,852	4.37%	10.91%
		I prefer dealing in cash	1,397	3.30%	8.23%
Supply Inefficiency	High Cost	It is expensive to have a bank account	2,479	5.85%	14.60%
		Charges and fees too high (e.g. opening fee, monthly fee)	527	1.24%	3.10%
		It costs too much to reach a bank	1,722	4.06%	10.14%
		Profit/interest is low	1,112	2.62%	6.55%
	Distance	Bank is far away	4,166	9.83%	24.54%
	Poor Service	Snobbish staff/not helpful	289	0.68%	1.70%
		Staff doesn't understand my needs	348	0.82%	2.05%
		Staff speaks in complicated terms	407	0.96%	2.40%
		Inconvenient hours of operation/Processing too long/Not user friendly	303	0.72%	1.78%
		Accessibility is poor for disabled people	527	1.24%	3.10%
		Banks are overcrowded/long queues	206	0.49%	1.21%
	No Trust	Don't trust banks or other financial institutes	790	1.86%	4.65%
		Too much corruption (e.g. bribes)	428	1.01%	2.52%
	Safety	Feel unsafe when leaving bank with money	267	0.63%	1.57%
		Scared or uneasy in a bank environment	309	0.73%	1.82%

[a]Note: Summary Data is derived from the unweighted dataset.

Table 4: Probit Regressions – Respondent Financial Status, Marginal Effects

VARIABLES	(1) Formally Banked	(2) Formally Banked	(3) Informally Banked	(4) Informally Banked	(5) Unbanked	(6) Unbanked
Individual Level Variables						
LnPerIncome	0.01***	0.01***	0.02***	0.01***	-0.01***	-0.01***
	(0.001)	(0.002)	(0.003)	(0.004)	(0.001)	(0.002)
Formality_documents	0.02***	0.02***	-0.02**	-0.01	-0.02***	-0.02***
	(0.002)	(0.002)	(0.005)	(0.006)	(0.002)	(0.002)
Completed Primary	0.01	0.02	0.02	0.04	-0.01	-0.02
	(0.011)	(0.012)	(0.017)	(0.021)	(0.011)	(0.012)
Completed Secondary	0.07***	0.09***	0.01	0.03	-0.07***	-0.09***
	(0.015)	(0.017)	(0.021)	(0.023)	(0.015)	(0.017)
Above Secondary	0.29***	0.32***	-0.18***	-0.17***	-0.29***	-0.32***
	(0.027)	(0.033)	(0.020)	(0.023)	(0.027)	(0.033)
Age/100	0.44***	0.41***	1.28***	1.30***	-0.44***	-0.41***
	(0.078)	(0.091)	(0.179)	(0.196)	(0.078)	(0.091)
Age/100_sq	-0.37***	-0.32***	-1.37***	-1.39***	0.37***	0.32***
	(0.088)	(0.098)	(0.175)	(0.207)	(0.088)	(0.098)
Female	0.00	0.00	0.02	0.02*	-0.00	-0.00
	(0.007)	(0.007)	(0.013)	(0.012)	(0.007)	(0.007)
Mobile	0.10***	0.11***	0.02	0.02	-0.10***	-0.11***
	(0.008)	(0.011)	(0.015)	(0.016)	(0.008)	(0.011)
TimetoStore	-0.00	0.00	0.00	0.00	0.00	-0.00
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Urban	0.02***	0.02	-0.01	-0.00	-0.02***	-0.02
	(0.007)	(0.010)	(0.019)	(0.020)	(0.007)	(0.010)
BankTrust	0.08***	0.08***	-0.07***	-0.07***	-0.08***	-0.08***
	(0.008)	(0.010)	(0.017)	(0.019)	(0.008)	(0.010)
RiskAversion	-0.02***	-0.02***	-0.00	-0.01	0.02***	0.02***
	(0.005)	(0.006)	(0.021)	(0.020)	(0.005)	(0.006)
FSKnow	0.02***	0.02***	-0.01	-0.01	-0.02***	-0.02***
	(0.002)	(0.002)	(0.005)	(0.005)	(0.002)	(0.002)
State Level Variables						
used_bribe		0.01		-0.05		-0.01
		(0.039)		(0.097)		(0.039)
societal_trust		-0.00		-0.02		0.00
		(0.034)		(0.069)		(0.034)
informality_no_income_taxes		-0.07		0.28***		0.07
		(0.038)		(0.080)		(0.038)
crime_level		-0.04		0.14		0.04
		(0.046)		(0.078)		(0.046)
GDP_per_capita_PPP		0.00		0.00		-0.00
		(0.000)		(0.000)		(0.000)
branchespermillion		0.01**		-0.02*		-0.01**
		(0.003)		(0.008)		(0.003)
atmspermillion		-0.01**		0.02**		0.01**
		(0.002)		(0.007)		(0.002)
mfospermillion		-0.00		-0.00		0.00
		(0.001)		(0.003)		(0.001)
Observations	21,109	21,109	21,109	21,109	21,109	21,109
State Controls	Yes		Yes		Yes	

State controls comprise of a series of state dummy variables.
Unbanked category includes those financially excluded and the informally banked.
Less than completed primary education is the omitted education category.
Robust standard errors, clustered by state in parentheses.
*** p<0.001, ** p<0.01, * p<0.05

Table 5: Reasons for being Unbanked – Multivariate Probit Results

VARIABLES	(1) Supply Constraints	(2) Supply Constraints	(3) Involuntary Exclusion	(4) Involuntary Exclusion	(5) Voluntary Exclusion	(6) Voluntary Exclusion
Individual Level Variables						
LnPerIncome	0.06***	0.05**	-0.03*	-0.05**	-0.00	-0.02
	(0.015)	(0.018)	(0.016)	(0.017)	(0.014)	(0.015)
Formality_documents	0.06**	0.07**	-0.04*	-0.02	0.05*	0.07**
	(0.022)	(0.024)	(0.015)	(0.019)	(0.025)	(0.025)
Completed Primary	0.09	0.17**	-0.05	-0.08	0.06	0.11
	(0.047)	(0.054)	(0.054)	(0.060)	(0.067)	(0.062)
Completed Secondary	0.07	0.18*	-0.06	-0.15*	0.09	0.18*
	(0.066)	(0.075)	(0.071)	(0.074)	(0.088)	(0.082)
Above Secondary	0.08	0.21*	-0.09	-0.14	0.10	0.22*
	(0.096)	(0.097)	(0.131)	(0.132)	(0.122)	(0.114)
Age/100	1.85***	1.76**	-0.44	0.15	0.92	1.04
	(0.517)	(0.550)	(0.665)	(0.632)	(0.736)	(0.721)
Age/100_sq	-2.19***	-2.05***	0.32	-0.31	-0.92	-1.11
	(0.558)	(0.581)	(0.750)	(0.720)	(0.870)	(0.879)
Female	-0.07*	-0.05	-0.01	-0.02	0.03	0.05
	(0.033)	(0.031)	(0.031)	(0.030)	(0.045)	(0.045)
Mobile	0.01	0.05	-0.04	-0.05	-0.03	-0.00
	(0.056)	(0.061)	(0.050)	(0.054)	(0.037)	(0.047)
urban	-0.21**	-0.29***	0.08	0.18**	-0.01	-0.04
	(0.081)	(0.062)	(0.057)	(0.060)	(0.068)	(0.069)
TimetoStore	0.00**	0.00*	-0.00	-0.00	-0.00	-0.00
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
BankTrust	0.03	0.01	0.10	0.13*	-0.02	-0.03
	(0.071)	(0.070)	(0.056)	(0.052)	(0.063)	(0.061)
RiskAversion	0.11	0.14	0.00	0.02	0.02	0.03
	(0.074)	(0.070)	(0.055)	(0.056)	(0.066)	(0.064)
FSKnow	-0.02	-0.02	-0.00	0.01	-0.03*	-0.04**
	(0.013)	(0.014)	(0.019)	(0.018)	(0.014)	(0.015)
State Level Variables						
used_bribe		0.10		-0.35		-0.15
		(0.383)		(0.256)		(0.352)
societal_trust		-0.17		-0.02		-0.25
		(0.300)		(0.245)		(0.241)
informality_no_income_taxes		-0.47		0.11		-0.68*
		(0.497)		(0.299)		(0.272)
crime_level		0.32		0.28		0.21
		(0.401)		(0.298)		(0.337)
GDP_per_capita_PPP		0.00		-0.00		-0.00
		(0.000)		(0.000)		(0.000)
branchespermillion		-0.01		-0.01		0.03
		(0.025)		(0.022)		(0.032)
atmspermillion		-0.01		0.01		-0.05
		(0.023)		(0.020)		(0.029)
mfospermillion		-0.02*		0.00		0.01
		(0.008)		(0.009)		(0.008)
Constant	-0.61***	-0.37	1.20***	1.02**	-1.04***	-0.49
	(0.113)	(0.461)	(0.114)	(0.311)	(0.135)	(0.357)
Observations	16,978	16,978	16,978	16,978	16,978	16,978
State Controls	Yes		Yes		Yes	

State controls comprise of a series of state dummy variables.
Less than completed primary education is the omitted education category.
Robust standard errors, clustered by state in parentheses.
*** p<0.001, ** p<0.01, * p<0.05

Table 6: Reasons for being Unbanked, Voluntary and Involuntary Exclusion – Multivariate Probit Results
Separate Multivariate Probit Regressions for Voluntary and Involuntary Constraints

	Involuntary Exclusion								Voluntary Exclusion			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Informality		Poverty		Financial Illiteracy		Low Education		Religion		Preference	
Individual Level Variables												
LnPerIncome	0.03	0.02	-0.06***	-0.06***	0.02	-0.00	-0.00	-0.03	0.07**	0.06**	-0.01	-0.03
	(0.019)	(0.020)	(0.014)	(0.015)	(0.017)	(0.017)	(0.022)	(0.020)	(0.025)	(0.019)	(0.015)	(0.016)
Formality_documents	-0.01	0.01	-0.03	-0.00	-0.02	-0.01	-0.05*	-0.04	-0.00	0.03	0.05*	0.06*
	(0.022)	(0.021)	(0.013)	(0.016)	(0.021)	(0.021)	(0.027)	(0.027)	(0.046)	(0.039)	(0.025)	(0.025)
Completed Primary	0.06	0.05	0.18**	0.13	0.02	0.03	-0.60***	-0.59***	0.00	0.01	0.07	0.12
	(0.069)	(0.073)	(0.066)	(0.071)	(0.058)	(0.055)	(0.107)	(0.103)	(0.104)	(0.084)	(0.068)	(0.064)
Completed Secondary	0.17**	0.16*	0.10	-0.02	0.03	0.06	-0.77***	-0.76***	0.16	0.16	0.09	0.19*
	(0.054)	(0.065)	(0.082)	(0.087)	(0.062)	(0.062)	(0.101)	(0.094)	(0.129)	(0.105)	(0.089)	(0.083)
Above Secondary	0.19	0.24*	0.08	-0.02	-0.04	0.04	-1.10***	-1.00***	0.13	0.19	0.09	0.21
	(0.114)	(0.110)	(0.112)	(0.117)	(0.113)	(0.111)	(0.181)	(0.179)	(0.196)	(0.173)	(0.131)	(0.122)
Age/100	1.08	1.41*	-0.91	-0.54	-0.12	0.05	0.60	0.81	0.60	0.63	0.95	1.05
	(0.703)	(0.681)	(0.542)	(0.500)	(0.449)	(0.426)	(0.621)	(0.647)	(1.322)	(1.215)	(0.791)	(0.768)
Age/100_sq	-1.07	-1.44	0.80	0.47	-0.22	-0.46	0.29	-0.04	-0.13	-0.30	-1.06	-1.19
	(0.840)	(0.807)	(0.660)	(0.619)	(0.505)	(0.497)	(0.601)	(0.630)	(1.401)	(1.264)	(0.927)	(0.929)
Female	-0.01	0.00	0.03	0.04	-0.05	-0.06	0.03	0.02	-0.03	0.01	0.03	0.04
	(0.039)	(0.039)	(0.052)	(0.049)	(0.044)	(0.041)	(0.057)	(0.055)	(0.091)	(0.091)	(0.044)	(0.043)
Mobile	-0.02	-0.02	0.04	0.01	0.02	0.03	-0.19**	-0.15*	0.06	0.04	-0.03	-0.01
	(0.061)	(0.056)	(0.054)	(0.058)	(0.051)	(0.047)	(0.060)	(0.060)	(0.096)	(0.094)	(0.035)	(0.047)
urban	-0.16**	-0.12*	0.06	0.20**	-0.03	-0.08	-0.16**	-0.10	0.12	0.10	-0.02	-0.04
	(0.061)	(0.061)	(0.063)	(0.068)	(0.060)	(0.064)	(0.063)	(0.052)	(0.096)	(0.099)	(0.072)	(0.070)
TimetoStore	0.00	0.00	0.00	0.00	0.00	-0.00	-0.01**	-0.01**	0.00	-0.00	-0.00	-0.00
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)
BankTrust	0.07	0.08	0.08	0.11*	-0.10*	-0.09	-0.02	0.02	0.16*	0.21**	-0.04	-0.05
	(0.073)	(0.068)	(0.053)	(0.055)	(0.041)	(0.045)	(0.095)	(0.102)	(0.080)	(0.075)	(0.063)	(0.061)
RiskAversion	0.05	0.07	-0.10*	-0.08	-0.01	0.03	-0.06	-0.05	-0.09	-0.01	0.03	0.03
	(0.070)	(0.071)	(0.039)	(0.043)	(0.051)	(0.050)	(0.066)	(0.067)	(0.087)	(0.080)	(0.065)	(0.064)
FSKnow	0.00	0.00	0.04**	0.04***	-0.06***	-0.06***	-0.10***	-0.10***	-0.04	-0.04	-0.03*	-0.04**
	(0.016)	(0.017)	(0.014)	(0.014)	(0.015)	(0.015)	(0.025)	(0.022)	(0.022)	(0.022)	(0.014)	(0.015)
State Level Variables												
used_bribe		-0.21		-0.00		-0.22		-0.70*		-0.13		-0.14
		(0.374)		(0.321)		(0.389)		(0.308)		(0.572)		(0.342)
societal_trust		-0.07		-0.01		-0.15		0.16		-0.22		-0.25
		(0.313)		(0.336)		(0.313)		(0.269)		(0.352)		(0.233)
informality_no_income_taxes		-0.55		0.84*		-0.66*		-0.76*		-0.37		-0.68*
		(0.298)		(0.381)		(0.273)		(0.330)		(0.460)		(0.266)
crime_level		-0.01		0.16		0.03		0.53		0.03		0.22
		(0.386)		(0.388)		(0.299)		(0.417)		(0.462)		(0.333)
GDP_per_capita_PPP		0.00		-0.00		0.00		-0.00		-0.00		0.00
		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
branchespermillion		0.03		0.01		-0.09**		0.02		0.04		0.03
		(0.035)		(0.027)		(0.028)		(0.028)		(0.039)		(0.031)
atmspermillion		-0.05		0.00		0.07**		-0.02		-0.05		-0.05
		(0.030)		(0.024)		(0.024)		(0.026)		(0.037)		(0.029)
mfospermillion		-0.01		-0.00		0.01		0.02		-0.00		0.01
		(0.009)		(0.011)		(0.007)		(0.014)		(0.020)		(0.008)
Constant	-1.31***	-0.85*	0.39***	-0.31	-0.03	0.23	-0.37	-0.22	-2.13***	-1.88***	-1.12***	-0.51
	(0.138)	(0.388)	(0.097)	(0.348)	(0.095)	(0.319)	(0.207)	(0.385)	(0.284)	(0.505)	(0.146)	(0.355)
Observations	16,978	16,978	16,978	16,978	16,978	16,978	16,978	16,978	16,978	16,978	16,978	16,978
State Controls	Yes		Yes		Yes		Yes		Yes		Yes	

State controls comprise of a series of state dummy variables.
Less than completed primary education is the omitted education category.
Robust standard errors, clustered by state in parentheses.
*** p<0.001, ** p<0.01, * p<0.05

Table 7: Reasons for being Unbanked, Supply Constraints – Multivariate Probit Results

	(1)	(2)	(3)	(4)	(6)	(7)	(8)	(9)	(11)	(12)
	Safety Concerns		High Cost		Distance		Poor Service		Low Trust	
Individual Level Variables										
LnPerIncome	-0.01	-0.03*	0.02	0.02	0.03*	0.01	-0.01	-0.02	0.03	0.01
	(0.013)	(0.013)	(0.015)	(0.016)	(0.014)	(0.016)	(0.016)	(0.019)	(0.017)	(0.020)
Formality_documents	0.02	0.03	-0.04*	-0.03	-0.01	-0.00	0.01	0.01	-0.02	-0.02
	(0.029)	(0.026)	(0.015)	(0.017)	(0.016)	(0.017)	(0.016)	(0.017)	(0.017)	(0.015)
Completed Primary	0.15	0.22**	0.07	0.11	0.10*	0.19***	0.05	0.07	0.09	0.20**
	(0.095)	(0.084)	(0.067)	(0.069)	(0.045)	(0.047)	(0.090)	(0.096)	(0.072)	(0.065)
Completed Secondary	0.17	0.27***	0.03	0.09	0.03	0.13	0.17	0.19*	0.03	0.19***
	(0.088)	(0.078)	(0.074)	(0.075)	(0.062)	(0.066)	(0.092)	(0.098)	(0.059)	(0.052)
Above Secondary	-0.20	-0.09	-0.51***	-0.44***	-0.46***	-0.39***	-0.09	-0.04	-0.35**	-0.22
	(0.211)	(0.198)	(0.129)	(0.123)	(0.099)	(0.093)	(0.100)	(0.111)	(0.130)	(0.124)
Age/100	2.14*	2.44*	0.54	0.60	0.67	0.85	1.12	1.43	0.71	1.07
	(1.038)	(1.033)	(0.534)	(0.561)	(0.655)	(0.650)	(0.736)	(0.743)	(0.742)	(0.675)
Age/100_sq	-2.53*	-2.78*	-0.68	-0.77	-1.11	-1.18	-1.23	-1.55	-0.78	-1.04
	(1.145)	(1.147)	(0.566)	(0.590)	(0.709)	(0.695)	(0.839)	(0.798)	(0.846)	(0.741)
Female	-0.13	-0.11	-0.07	-0.05	-0.09**	-0.08*	-0.09	-0.07	-0.09	-0.07
	(0.084)	(0.081)	(0.038)	(0.036)	(0.033)	(0.031)	(0.046)	(0.043)	(0.045)	(0.046)
Mobile	-0.07	-0.08	-0.02	-0.00	-0.14**	-0.08	0.02	0.01	-0.12	-0.04
	(0.070)	(0.064)	(0.060)	(0.062)	(0.054)	(0.058)	(0.067)	(0.058)	(0.065)	(0.063)
urban	-0.15	-0.15	-0.03	-0.09	-0.52***	-0.52***	-0.08	-0.05	-0.06	-0.10
	(0.103)	(0.096)	(0.074)	(0.052)	(0.067)	(0.059)	(0.075)	(0.070)	(0.081)	(0.074)
TimetoStore	0.00	-0.00	0.00*	0.00	0.00**	0.00**	-0.00	-0.00	-0.00	-0.00
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
BankTrust	-0.02	-0.02	-0.10	-0.11	-0.09	-0.09	-0.08	-0.06	-0.14	-0.14
	(0.087)	(0.075)	(0.065)	(0.063)	(0.072)	(0.069)	(0.072)	(0.064)	(0.083)	(0.083)
RiskAversion	0.14	0.14*	0.06	0.08	0.13	0.15*	0.05	0.09	0.26***	0.24***
	(0.076)	(0.070)	(0.058)	(0.052)	(0.078)	(0.076)	(0.068)	(0.061)	(0.063)	(0.060)
FSKnow	-0.09***	-0.09***	-0.07***	-0.06***	-0.06***	-0.04**	-0.07***	-0.07***	-0.04**	-0.03**
	(0.018)	(0.018)	(0.012)	(0.013)	(0.015)	(0.016)	(0.014)	(0.013)	(0.012)	(0.010)
State Level Variables										
used_bribe		0.03		0.47		-0.25		-0.24		0.01
		(0.302)		(0.294)		(0.252)		(0.522)		(0.290)
societal_trust		-0.31		-0.28		0.11		-0.09		-0.12
		(0.331)		(0.265)		(0.232)		(0.346)		(0.395)
informality_no_income_taxes		0.17		-0.40		-0.27		-0.32		-0.17
		(0.243)		(0.379)		(0.301)		(0.387)		(0.292)
crime_level		-0.25		-0.00		0.93***		-0.13		0.01
		(0.278)		(0.303)		(0.279)		(0.409)		(0.336)
GDP_per_capita_PPP		-0.00		0.00		0.00*		0.00		-0.00
		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
branchespermillion		0.06		-0.01		-0.04		0.08		0.09*
		(0.039)		(0.023)		(0.024)		(0.042)		(0.036)
atmspermillion		-0.06		-0.01		0.03		-0.09*		-0.09**
		(0.035)		(0.021)		(0.021)		(0.038)		(0.029)
mfospermillion		-0.02*		-0.01*		-0.01		-0.02*		-0.01
		(0.010)		(0.007)		(0.008)		(0.009)		(0.009)
Constant	-1.94***	-1.84***	-0.71***	-0.27	-0.71***	-1.03**	-1.42***	-1.10*	-1.55***	-1.66***
	(0.201)	(0.408)	(0.150)	(0.445)	(0.141)	(0.324)	(0.157)	(0.435)	(0.197)	(0.328)
Observations	21,109	21,109	21,109	21,109	21,109	21,109	21,109	21,109	21,109	21,109
State Controls	Yes		Yes		Yes		Yes		Yes	

State controls comprise of a series of state dummy variables.
Less than completed primary education is the omitted education category.
Robust standard errors, clustered by state in parentheses.
*** p<0.001, ** p<0.01, * p<0.05

Table 8: First Stage Regression, Ordinary Least Squares
Dependent Variable: Number of Formal Documents

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Type of Refuse Collection			Time to Water Source			Both Instruments		
Individual Level Variables									
Refuse_disposal	-0.13*	-0.13*	-0.14***				-0.13*	-0.13*	-0.14***
	(0.049)	(0.047)	(0.032)				(0.049)	(0.047)	(0.031)
Time_water				-0.09***	-0.08***	-0.08***	-0.09***	-0.08***	-0.08**
				(0.018)	(0.020)	(0.023)	(0.019)	(0.019)	(0.023)
LnPerIncome	0.10***	0.10***	0.10***	0.10***	0.10***	0.10***	0.10***	0.10***	0.10***
	(0.015)	(0.014)	(0.009)	(0.015)	(0.015)	(0.010)	(0.015)	(0.014)	(0.010)
Completed Primary	0.15***	0.14**	0.14***	0.15***	0.15**	0.14***	0.15***	0.15***	0.14***
	(0.038)	(0.041)	(0.035)	(0.039)	(0.043)	(0.036)	(0.037)	(0.041)	(0.035)
Completed Secondary	0.26***	0.24**	0.24***	0.27***	0.26**	0.24***	0.26***	0.25**	0.24***
	(0.064)	(0.070)	(0.050)	(0.065)	(0.073)	(0.050)	(0.063)	(0.071)	(0.050)
Above Secondary	0.97***	0.99***	0.95***	0.98***	1.01***	0.96***	0.96***	0.99***	0.94***
	(0.096)	(0.108)	(0.080)	(0.100)	(0.111)	(0.081)	(0.097)	(0.110)	(0.080)
Age/100	5.07***	5.29***	4.96***	5.07***	5.30***	4.96***	5.08***	5.29***	4.96***
	(0.789)	(0.750)	(0.542)	(0.801)	(0.765)	(0.547)	(0.796)	(0.754)	(0.542)
Age/100_sq	-2.67**	-2.97**	-2.49***	-2.66**	-2.97**	-2.48***	-2.67**	-2.96**	-2.49***
	(0.878)	(0.860)	(0.626)	(0.892)	(0.878)	(0.632)	(0.885)	(0.865)	(0.625)
Female	-0.53***	-0.53***	-0.53***	-0.53***	-0.53***	-0.53***	-0.53***	-0.54***	-0.53***
	(0.047)	(0.047)	(0.030)	(0.047)	(0.047)	(0.030)	(0.047)	(0.047)	(0.030)
Mobile	0.37***	0.35***	0.33***	0.37***	0.34***	0.33***	0.36***	0.34***	0.32***
	(0.039)	(0.043)	(0.036)	(0.038)	(0.042)	(0.036)	(0.039)	(0.042)	(0.036)
urban	0.08	0.05	0.07	0.08	0.05	0.07	0.06	0.03	0.05
	(0.051)	(0.060)	(0.078)	(0.051)	(0.060)	(0.079)	(0.049)	(0.060)	(0.079)
TimetoStore	-0.00	-0.00	0.00	0.00	0.00	0.00	-0.00	0.00	0.00
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
BankTrust	0.09	0.10*	0.09*	0.09	0.09	0.09*	0.09	0.10	0.09*
	(0.048)	(0.049)	(0.036)	(0.047)	(0.049)	(0.037)	(0.047)	(0.048)	(0.036)
RiskAversion	-0.16*	-0.15*	-0.17***	-0.15*	-0.14	-0.17***	-0.15*	-0.15*	-0.17***
	(0.067)	(0.070)	(0.044)	(0.067)	(0.070)	(0.043)	(0.067)	(0.070)	(0.043)
FSKnow	0.13***	0.12***	0.12***	0.13***	0.12***	0.12***	0.12***	0.12***	0.12***
	(0.011)	(0.013)	(0.010)	(0.012)	(0.013)	(0.010)	(0.012)	(0.013)	(0.010)
State Level Variables									
used_bribe		-0.23	0.19		-0.24	0.17		-0.22	0.21
		(0.214)	(0.143)		(0.231)	(0.151)		(0.215)	(0.143)
societal_trust		-0.04	-2.50***		-0.04	-2.34***		-0.04	-2.39***
		(0.183)	(0.077)		(0.182)	(0.090)		(0.182)	(0.090)
informality_no_income_taxes		-0.95***	4.12***		-0.88***	4.31***		-0.92***	4.23***
		(0.226)	(0.101)		(0.234)	(0.106)		(0.224)	(0.103)
crime_level		-0.18	-3.38***		-0.12	-3.06***		-0.15	-3.27***
		(0.303)	(0.168)		(0.313)	(0.192)		(0.307)	(0.173)
GDP_per_capita_PPP		0.00	0.00*		0.00	0.00		0.00	0.00
		(0.000)	(0.000)		(0.000)	(0.000)		(0.000)	(0.000)
branchespermillion		-0.00	-0.13***		0.00	-0.18***		-0.00	-0.15***
		(0.026)	(0.018)		(0.025)	(0.020)		(0.026)	(0.020)
atmspermillion		0.01	0.20***		0.00	0.25***		0.01	0.23***
		(0.022)	(0.019)		(0.021)	(0.020)		(0.022)	(0.020)
mfospermillion		0.01	-0.10***		0.01	-0.10***		0.01	-0.10***
		(0.008)	(0.002)		(0.007)	(0.002)		(0.008)	(0.002)
Constant	-0.53*	0.12	-0.57***	-0.66***	-0.13	-0.95***	-0.32	0.25	-0.46**
	(0.220)	(0.273)	(0.168)	(0.178)	(0.236)	(0.143)	(0.223)	(0.284)	(0.170)
Observations	21,109	21,109	21,109	21,109	21,109	21,109	21,109	21,109	21,109
State Controls	Yes			Yes			Yes		
LGA Controls			Yes			Yes			Yes
R-squared	0.36	0.34	0.40	0.36	0.34	0.40	0.36	0.34	0.40
Adj. R-squared	0.35	0.34	0.39	0.35	0.34	0.38	0.36	0.34	0.39
Angrist-Pischke F-Test							15.84	10.62	15.13

State controls comprise of a series of state dummy variables.
Local government area controls comprise of a series of 586 dummy variables.
Robust standard errors, clustered by state (district in columns 3, 6 and 9) in parentheses.
*** p<0.001, ** p<0.01, * p<0.05

Table 9: Formally Banked and No. of Formal Documents, Probit and Instrumental Variables Estimates

	(1)	(2)	(3)	(4)
	Probit		Both IVs	
Individual Level Variables				
Formality_documents	0.17***	0.16***	0.54**	0.54**
	(0.016)	(0.016)	(0.174)	(0.175)
LnPerIncome_p_plus	0.10***	0.09***	0.04	0.04
	(0.009)	(0.010)	(0.042)	(0.040)
Completed Primary	0.10	0.16	0.02	0.07
	(0.087)	(0.087)	(0.103)	(0.110)
Completed Secondary	0.49***	0.57***	0.29	0.36
	(0.090)	(0.087)	(0.192)	(0.203)
Above Secondary	1.27***	1.31***	0.64	0.65
	(0.099)	(0.100)	(0.496)	(0.520)
Age/100	3.53***	3.04***	0.86	0.30
	(0.625)	(0.647)	(1.928)	(1.950)
Age/100_sq	-2.99***	-2.40***	-1.40	-0.73
	(0.712)	(0.715)	(1.342)	(1.326)
Female	0.01	0.03	0.22*	0.24*
	(0.054)	(0.051)	(0.111)	(0.106)
Mobile	0.73***	0.76***	0.45*	0.48*
	(0.057)	(0.060)	(0.224)	(0.234)
urban	0.19***	0.13*	0.11	0.08
	(0.048)	(0.065)	(0.067)	(0.090)
TimetoStore	-0.00	0.00	0.00	0.00
	(0.002)	(0.002)	(0.002)	(0.002)
BankTrust	0.57***	0.51***	0.44**	0.38**
	(0.050)	(0.047)	(0.152)	(0.143)
RiskAversion	-0.14**	-0.16**	-0.05	-0.07
	(0.047)	(0.053)	(0.086)	(0.096)
FSKnow	0.17***	0.17***	0.09	0.09
	(0.018)	(0.016)	(0.069)	(0.067)
State Level Variables				
used_bribe		0.06		0.15
		(0.292)		(0.258)
societal_trust		-0.03		-0.00
		(0.253)		(0.227)
informality_no_income_taxes		-0.50		-0.04
		(0.284)		(0.494)
crime_level		-0.27		-0.15
		(0.341)		(0.242)
GDP_per_capita_PPP		0.00		0.00
		(0.000)		(0.000)
branchespermillion		0.06*		0.04
		(0.023)		(0.024)
atmspermillion		-0.05*		-0.04
		(0.020)		(0.022)
mfospermillion		-0.00		-0.01
		(0.010)		(0.009)
Constant	-4.27***	-3.87***	-3.15**	-3.03**
	(0.194)	(0.326)	(1.152)	(1.019)
Observations	21,109	21,109	21,109	21,109
State Controls	Yes		Yes	
Hansen J P-Value			N/A	0.5879

State controls comprise of a series of state dummy variables
Robust standard errors, clustered by state in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table 10: Marginal Effects, Clustered by State at Alternative Levels of Formal Documents.
IV: Categorical Version of Instruments

	With State Level variables		With State Controls	
Marginal Effects at:	Probit	IV Probit	Probit	IV Probit
mean	0.021***	0.080	0.021***	0.077
0	0.021***	0.055***	0.021***	0.055***
1	0.022***	0.092	0.022***	0.089
2	0.024***	0.131	0.024***	0.124
3	0.025***	0.160	0.025***	0.152
4	0.027***	0.168*	0.027***	0.161
5	0.028***	0.147***	0.028***	0.146***
6	0.030***	0.105	0.030***	0.110
7	0.031***	0.060	0.031***	0.067
8	0.033***	0.027	0.033***	0.033
9	0.034***	0.009	0.034***	0.013
10	0.036***	0.002	0.035***	0.004
11	0.037***	0.001	0.036***	0.001
12	0.038***	0	0.037***	0
13	0.035***	0	0.038***	0
14	0.038***	0	0.038***	0
15	0.039***	0	0.038***	0

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Institute for International Integration Studies
The Sutherland Centre, Trinity College Dublin, Dublin 2, Ireland