University of Leicester

DEPARTMENT OF ECONOMICS

Loan Defaults in Africa*

Svetlana Andrianova, University of Leicester, UK
Badi H. Baltagi, Syracuse University, US, and University of Leicester, UK
Panicos O. Demetriades, University of Leicester, UK

Working Paper No. 11/36
August 2011

Loan Defaults in Africa*

Svetlana Andrianova,[†] *Badi H. Baltagi,*[‡] *and Panicos O. Demetriades*[§]

August 4, 2011

Abstract. African financial deepening is beset by a high rate of loan defaults, which encourages banks to hold liquid assets instead of lending. We put forward a novel theoretical model that captures the salient features of African credit markets which shows that equilibrium with high loan defaults and low lending can arise when contract enforcement institutions are weak, investment opportunities are relatively scarce and information imperfections abound. We provide evidence using a panel of 110 banks from 29 African countries which corroborates our theoretical predictions.

Keywords: Financial development, Africa

JEL: G21, O16

---

*We acknowledge financial support from the Economic and Social Research Council (Award reference RES-000-22-2774). We would like to thank the participants of the Money, Macro and Finance Research Group workshop in Birmingham and the "Banking and Finance Post-Crisis" Conference at the Leicester University Institute of Finance, and seminar participants at the University of Leicester. Special thanks are due to Charles Calomiris, Alex Trew, and Vincenzo Denicolò.

† University of Leicester, U.K.

‡ Syracuse University, U.S., and University of Leicester, U.K.

§ Corresponding author. Address for correspondence: Professor Panicos O. Demetriades, Department of Economics, University of Leicester, University Road, Leicester LE1 7RH, UK. E-mail: pd28@le.ac.uk. Tel:+ 44 (0) 116 252 2835. Fax: + 44 (0) 116 252 2908.

1 Introduction

Sub-Saharan Africa remains one of the most financially under-developed regions in the world (Honohan and Beck 2007). Banks in Africa complain that there is a lack of creditworthy borrowers while at the same time households and firms find finance as a major constraint in their activities. Recent research has shown that banks are deterred from lending by a very high rate of loan defaults (Andrianova, Baltagi, Demetriades and Fielding 2011, Demetriades and Fielding 2010). Understanding the determinants of loan default rates in Africa seems, therefore, to hold the key to overcoming the obstacles to financial development in Africa. This paper makes a first step in this direction by providing both theory and evidence that shed new light on the factors behind high loan defaults in Africa.

The starting point in our theoretical analysis is a model of bank lending with both adverse selection and moral hazard, resulting from three different types of borrowers: honest, who always repay, dishonest, who always default, and opportunistic who can choose to repay or default, depending on their expected payoff. We assume that some banks have access to an imperfect screening technology, which they can choose to use in order to reduce adverse selection. We also assume that the output of the project is a form of collateral, along the lines suggested by Rousseau (1998). In our case, however, the fraction of the loan that is recovered is determined by the degree of contract enforcement; this is a key policy parameter which determines the type of equilibrium that obtains. Such a setting captures the stylised facts of African credit markets rather well, resulting in equilibria in which a high rate of loan defaults co-exists with banks' choosing to lend a fraction of their available resources.[1]

The key predictions of the theory are tested using a panel data set comprising 110 banks from 29 countries over the period 1998-2008. The empirical model, which is in-

[1]With few notable exceptions (Shubik 1973, Stiglitz and Weiss 1981, Rousseau 1998), loan defaults in economic models are not an equilibrium phenomenon. For an excellent overview of the neglected role of loan defaults in economics see Goodhart and Tsomokos (2011).

formed by the theory, assumes that loan defaults are determined by bank characteristics, such as age and size, and country-wide variables purporting to capture the quality of contract enforcement and the availability of investment opportunities. The model is estimated using the Papke and Wooldridge (1996) fractional logit regression which handles fractional response variables based on quasi-likelihood methods. The results of these estimations, which are consistent with our theoretical predictions, provide an empirical handle on the relative influence of different types of institutions and economic growth on the default rate.

The paper is organised as follows. Section 2 sets out the theoretical model and derives its key predictions. Section 3 presents our empirical methods and data. Section 4 presents and discusses the empirical results. Section 5 summarises and concludes.

2 Theory

The starting point of our model is the "linear city".[2] There are three types of entrepreneurs (borrowers) who seek a loan of 1 monetary unit in order to undertake an investment project: "honest" (in proportion α), "dishonest" (in proportion β) and "opportunistic" (in proportion γ with $\alpha + \beta + \gamma = 1$). The honest type has an investment project with rate of return R and this borrower always repays the loan. The opportunistic type has the same investment project (with rate of return R) but can choose whether to repay or default on his loan. The dishonest type has a project with rate of return 0, this borrower will always default on his loan. The borrower's type is private information, the proportions α, β and γ are publicly known. All borrowers are uniformly distributed on a unitary interval with distribution density 1.[3] Applying for a loan is costly for a borrower due to the transportation cost of t per unit of distance between the borrower

[2]The basic model is due to Hotelling (1929), but it is extended here along the lines which were explored in Andrianova et al. (2011). The key difference between the current model and these two previous papers is that it incorporates asymmetric information on both sides of the transaction.

[3]Geographical distance captures difference in individual tastes for the offered loan contract.

and the lender. Each borrower can apply for a loan to at most one lender.

Two lenders (banks) are available and located at the opposite ends of the interval (bank A is at 0, bank B is at 1). The lenders compete for loan contracts: bank i sets its loan interest rate r_i to maximise its expected payoff ($i = \{A, B\}$). Each lender has sufficient funds to approve all loan applications.[4] The banks differ, in principle, in their ability to gather and use information about individual borrower's characteristics. Either lender may be of one of the two types: "competent" type has a screening technology, while "incompetent" has no ability to screen its borrowers. The lender is competent with probability κ, otherwise with probability $1-\kappa$ it is incompetent; κ is common knowledge, while the type of the lender is its private information. The screening technology, if used, allows to obtain with probability σ a negative bundled signal about an individual borrower who is either an opportunist with R-return project or a dishonest borrower with a zero-return project. With probability $1 - \sigma$ the screening fails to find any information about the borrower/project. With slight abuse of notation, κ-type bank chooses whether to screen or not to screen its borrowers, and then whether to refuse (or not) applications of borrowers with a negative screening signal. All lenders have access to a "safe" asset which has the rate of return r_0 ($0 < r_0 < R$).

The loan contract enforcement is imperfect. Loan default when investment returned R is remedied with probability λ by an imposition of a monetary penalty of $1 + R$ on the defaulted borrower, of which the lender gets $1 + r_i$. Thus, the lender, if compensated, is "made good" according to the contractual terms of his loan contract. The difference between the borrower's penalty and lender's compensation is the enforcement cost and is borne by the defaulted borrower. With probability $1 - \lambda$, such default is unenforceable (penalty and compensation are zero). When the borrower has zero return, no enforcement of the loan contract is possible. Investment return is ex post observable and non-falsifiable. All players are risk-neutral.

The timing of events is as follows:

[4] This assumption accords well with the findings of Honohan and Beck (2007).

(1) Bank i ($i = A, B$) sets its lending rate r_i.

(2) Each borrower chooses the bank in which to apply for a loan of 1 monetary unit.

(3) Facing the demand for loans, D_i, κ-type bank i chooses whether to screen or not all of its loan applications.

(4) Each bank chooses which applications to approve and which to decline.

(5) Honest and opportunistic borrowers with an approved loan invest.

(6) An honest borrower repays, a dishonest borrower defaults, an opportunistic borrower chooses whether to repay or default.

(7) If any of its loans are defaulted on, the bank seeks compensation.

(8) Payoffs are realized.

Let $q \in \{0,\ 1\}$ denote an opportunistic borrower's decision to repay ($q = 1$) or default ($q = 0$); let $\xi \in \{0,\ 1\}$ be κ-type bank's decision to screen ($\xi = 1$ is "screen') and let $p_j \in \{0,\ 1\}$ (where $j = \{\kappa,\ 1-\kappa\}$) be j bank's decision to approve all its applications when there is no screening.

Note that if the proportion of dishonest borrowers is negligible, the competent type of the bank will find it unprofitable to use the screening technology. Also, the larger the transportation cost the lower is the incentive of the dishonest borrower to apply for a loan. In what follows we will therefore assume that the proportion of borrowers is non-negligible and the transportation cost is not too high. Solving the model backwards for pooling equilibria, we obtain the following result:

Proposition 1 *There exist $\underline{\lambda}$, $\bar{\lambda}$, $\bar{\beta}$ and $\bar{t}$, so that for $\beta \geq \bar{\beta}$ and $t \leq \bar{t}$ the unique equilibrium of the game is:*

(1) the low default equilibrium (LDE) with $q = \xi = p_{1-\kappa} = 1$ if $\lambda \geq \bar{\lambda}$,

(2) the high default equilibrium (HDE) with $q = 0$, $\xi = p_{1-\kappa} = 1$ if $\underline{\lambda} \leq \lambda < \bar{\lambda}$,

(3) the no lending equilibrium (NLE) with $q = \xi = p_\kappa = p_{1-\kappa} = 0$ *if* $\lambda < \underline{\lambda}$.

Intuitively, NLE obtains when the contract enforcement is very weak. In such a situation, widespread default by opportunistic and dishonest borrowers makes lending unprofitable for any screening technology. LDE exists in the presence of a non-negligible proportion of dishonest borrowers when the contract enforcement is sufficiently good. And in the intermediate range of the contract enforcement parameter values, HDE obtains because the opportunistic borrowers have no incentive to repay (given relatively weak enforcement) but the banks nevertheless find it profitable to screen and lend to all those borrowers who do not appear to have a negative signal (given that enforcement is not too weak and defaults are compensated often enough to make lending profitable).

In this model, economic "growth" can be measured by the increase in $(1-\beta)$ because this proportion represents all borrowers who have the high return investment opportunity. Additionally, in HDE the default probability is measured by $(1-\alpha)[1-\kappa\sigma]$, which is decreasing in α, κ and σ. It is then easy to see that with higher economic growth the probability of default will fall as long as the increase in growth does not exclusively benefit opportunistic borrowers. In LDE, because opportunists do not default and consequently the default probability is $\beta(1-\kappa\sigma)$, economic growth will unambiguously reduce the default rate.

3 Methodology and Data

The first aim of the empirical model is to provide a framework in which the main predictions of the theory can be tested. Specifically, we aim to test whether (a) better contract enforcement and (b) an improvement in borrower quality reduce the probability of loan defaults. The second, equally important, aim of the empirical analysis is to shed light on what kind of policies can be used to contain loan defaults in Africa. Below we put forward a plausible empirical model that can address both aims.

It is reasonable to assume that the probability of loan default is a function of bank characteristics, including bank age and bank assets. We assume it could also depend on contract enforcement and other governance indicators capturing the institutional aspects of African credit markets, particularly those encapsulating the degree of contract enforcement. The model includes various interaction terms between the bank characteristics and the governance indicators in order to examine the effects of governance on different types of bank.

Formally, we assume that the proportion of loan defaults for bank i in country j at time t, denoted by D_{ijt}, is determined by the following relationship:

$$D_{ijt} = F(N_{ijt}, N_{ijt}^2, GS_{ijt}, LA_{ijt}, RoL_{jt}, Corr_{jt}, REG_{jt}, G_{jt}^Y, X_{ijt})$$

where N is the number of years a bank has been in operation, GS is the government's share in the bank, LA is the logarithm of total assets (expressed in million US dollars), RoL is rule of law, $Corr$ is control of corruption, REG is regulation, and GY is the growth rate of GDP. X is a vector of other controls that includes geographical dummies and interaction terms while subscripts i, j and t stand for bank, country and year, respectively.

If F is linear it can be estimated by OLS. However, there is no guarantee that the predicted proportions from OLS will lie between 0 and 1. Also OLS implies a constant marginal effect of the regressor on the probability of loan default. This may not be plausible, say for the effect of a one year increase in the age of the bank, as it may be different for a young upstart bank versus an old and well established bank. Despite these weaknesses, OLS remains a useful benchmark not least because the estimated parameters are easy to interpret. We therefore present OLS estimates of the model as the starting point in our empirical investigation. In all regressions we include year dummies to account for time varying macro shocks or common factors that affect all banks in all of our regressions. We find these time dummies statistically significant. As OLS does not account for the heterogeneity across banks, we also report results using a Random Effects (RE) estimator. Not accounting for such heterogeneity can lead to biased standard errors for the OLS estimates, wrong t-statistics and misleading inference.[5]

Like OLS, the RE estimator ignores the fact that the dependent variable is a fraction. We therefore apply the Papke and Wooldridge (1996) fractional logit regression which handles fractional response variables based on quasi-likelihood methods. Papke and Wooldridge (1996) propose modelling the conditional mean of the dependent variable as a logistic function. This ensures that the predicted value of the dependent variable lies in the interval (0, 1). It is also well defined even if the dependent variable takes the values 0 or 1 with positive probability (Gourieroux, Monfort and Trognon 1984). Maximizing the Bernoulli log-likelihood function yields the quasi-MLE (QMLE) which is consistent and asymptotically normal Gourieroux et al. (1984). McCullagh and Nelder (1989) propose the generalized linear model (GLM) approach to this problem. We apply the Logit QMLE in Stata using the GLM command with the Bernoulli Binary family function and the link function indicating the logistic distribution. To check robustness we also report results using the complementary log-log link function. The latter, unlike Logit, is asymmetric around zero and as such can account for skewness arising from a large number of zeros or ones in the dataset.

The bank data are collected from Bank Scope and consist of 110 banks in 29 countries observed over the period 1998–2008. We measure the default rate as the ratio of impaired loans to total loans. The governance indicators are from the World Bank Governance Database. GDP data and exchange rates are from World Development Indicators.

Summary statistics and correlations can be found in Tables 1 and 2. The mean default rate is just under 11.0 per cent and ranges from near 0 to 81 per cent. Bank age varies from a young 3 to a mature 147 years old and has a mean of just under 37 years. The government share varies from 0 to 100 and has a mean of 7.9 per cent. All the governance indicators we utilise have negative means and their standard deviation is

[5] Although Fixed Effects estimation accounts for the possible correlation of the random bank effects with the regressors, it wipes out important variables of interest such as bank age and is therefore not reported.

below 1; not surprisingly these compare unfavourably with a worldwide mean of 0 and a unit standard deviation. The mean growth rate, however, is a healthy 2.2 per cent per annum, but varies considerably from -16.4 per cent to 25.4 per cent.

The default rate is negatively correlated with bank age, positively correlated with government share and negatively correlated with the logarithm of assets. All these correlations of the default rate with individual bank characteristics are plausible, suggesting that as banks get bigger and older they acquire more information capital and are able to better control loan defaults. Government ownership is positively associated with loan defaults, indicating that politically determined priorities or indeed corruption result in a greater proportion of impaired loans. The default rate is negatively correlated with all the governance indicators, indicating that better governance results in lower loan defaults. Moreover, it is negatively correlated with the GDP growth rate, suggesting that as economic opportunities improve, loan defaults decline. All these, although plausible, are, however, unconditional correlations. The next section provides a more rigorous empirical analysis of the determinants of loan defaults.

4 Empirical Results

The empirical models include interactions between GS and Corr on the one hand and RoL and N on the other.[6] Additional controls include a set of time dummies and a North African dummy (which kicks in if a bank is located in Algeria, Egypt, Libya, Morocco or Tunisia).[7] We also report results by excluding the North African banks altogether.

[6]Other interactions were also attempted but were found insignificant and were, therefore, dropped.

[7]We also tried an 'offshore' dummy for Mauritius and the Seychelles but this was insignificant and therefore dropped.

OLS and RE Results

Table 3 presents an initial set of empirical estimates using OLS and RE. Where there are differences between OLS and RE, we attach more weight to the RE estimates because the Breusch-Pagan test suggests that the random effects are highly significant.

In both OLS and RE estimates, the North Africa dummy is positive and highly significant, while bank age appears to have an inverse U-shape effect on loan defaults as the level term is positive and significant while the squared term is negative and also significant. These estimates are consistent with a 'learning-by-doing' lending technology, with the turning point being around 1.5 years old, which seems plausible. *Logarithm of assets* enters with a negative and significant coefficient, albeit only at the 10% level in the OLS estimates, suggesting that bigger is better when it comes to reducing loan defaults. *Government share* is significant at the 10% level according to the OLS estimates but insignificant in the RE estimates. Its interaction with control of corruption is negative, although significant only in the OLS estimates, indicating that the effects of government ownership vary inversely with the degree of corruption. *Rule of law* is negative and significant (at the 1.0 per cent level in the RE estimates). Its effects appear to be tempered by bank age, as the interaction term between *Rule of law* and *Bank age* is positive and significant at the 1.0 per cent level. In both sets of estimates *Control of corruption* is negative and significant while *Regulatory quality* is, surprisingly, positive and highly significant. Finally, *Growth rate* is negative but is only significant in the OLS regressions.

When we exclude banks located in North Africa the results remain qualitatively very similar, although there are two exceptions. *Logarithm of assets* is no longer significant in the RE estimates, suggesting that within Sub-Saharan Africa, bigger banks are not able to better prevent loan defaults. The other key difference is that *Government share* is now positive and significant in both the RE and OLS estimates. Thus, it appears that government ownership of banks in Sub-Saharan Africa is associated with more loan defaults, although this effect is tempered when control of corruption is positive (i.e.

above the world mean).

GLM Results

Table 4 provides the GLM estimates of the same models, on which we attach more weight than the OLS and RE estimates. It should be noted that because of the non-linearity of the link function, the parameter estimates are not comparable to those presented in Table 3. Furthermore, the coefficients do not represent the marginal effect of a variable.

Qualitatively speaking the results in Table 4 are similar to those presented in Table 3 being, in fact, somewhat closer to those obtained with OLS than RE. To start with, the estimates for all countries suggest that *Bank age* has a positive effect while *Bank age squared* has a negative effect that is significant—both these effects are significant at the 1% level. *Logarithm of assets* has a negative effect that is marginally significant as was the case with both OLS and RE. *Government share* enters with a negative coefficient that is, however, far from being significant at conventional levels. This can be contrasted with a positive coefficient in Table 3, significant at the 10% level in the OLS case. Both *Rule of law* and *Control of corruption* have negative and significant effects, with the level of significance being 1% for both—a qualitatively similar result to that shown in Table 3. The same is true of *Regulatory quality* that once again enters with a positive coefficient that is significant at the 1% level. The *Growth rate* has a negative effect that is highly significant as was the case using OLS; in contrast the corresponding estimate by RE is not significant. The North Africa dummy is positive and highly significant as was the case with both OLS and RE. The interaction between *Government share* and *Control of corruption* is now negative and highly significant. The interaction between *Rule of law* and *Bank age* remains positive and highly significant.

The corresponding results using the complementary log-log link function are nearly identical to those obtained with the Logistic, indicating robustness to any skewness. Moreover, the time dummies are jointly significant and the diagnostics are satisfactory.

The estimates for Sub-Saharan Africa are qualitatively similar to those for all countries with three relatively minor exceptions. The first is the somewhat lower coefficients on *Bank age* and *Bank age squared* (in absolute value), indicating perhaps somewhat less rapid 'learning by doing'. The second is the coefficient of *Logarithm of assets* which loses significance and is about a third smaller in absolute terms. The third and final exception is that *Government share* remains insignificant, although it now has a positive coefficient.

The results presented in Table 4 confirm that the (lack of) economic growth is one of the most important causes of the high loan default rate, providing support to the adverse selection hypothesis. Moreover, they suggest that *Control of corruption* and *Rule of law* can help check the default rate, although their effect differs depending on a bank's age and ownership. Government owned banks stand to benefit most when control of corruption is above the world average while younger banks stand to benefit most from improvements in the rule of law.

The only surprising result is that *Regulatory quality* has a positive marginal effect on the default rate. This effect is of course *ceteris paribus*. As we are controlling for *Rule of law* and *Corruption*, and given the positive correlation between the three indicators, it could be argued that marginal improvements in regulatory quality, if not accompanied by improvements in *Rule of law* and *Corruption*, can only represent an increased burden on doing business i.e. 'red tape'.

5 Concluding Remarks

It is now well known that African financial deepening is plagued by a high rate of loan defaults, which deters banks from lending and encourages them to hold liquid domestic or foreign assets instead (Andrianova et al. 2011, Demetriades and Fielding 2010). A better understanding of the underlying causes of loan defaults, therefore, holds the key to addressing financial under-development in Africa. To this end, we have put forward a model of bank lending that captures some of the salient features of

African credit markets, including the lack of collateral, weak contract enforcement and severe information imperfections. In our model, loan defaults arise as an equilibrium phenomenon, in contrast to nearly all previous theoretical literature on credit markets. The incidence of loan defaults has been shown to depend inversely on the effectiveness of contract enforcement and the availability of investment opportunities.

It has been shown that weak contract enforcement can combine with lack of economic opportunities and rising adverse selection to deter banks from lending altogether. At the opposite end of the spectrum, when contract enforcement is reasonably good—albeit imperfect—banks will choose to lend, even if their screening technologies are unable to detect all dishonest borrowers, who will always default on their loans. We have also shown that there exists an intermediate equilibrium with high loan defaults. In this equilibrium both dishonest and opportunistic borrowers default on their loans. Banks, nevertheless, find it profitable to lend. Thus, high loan defaults—a stylised fact of African credit markets—can co-exist with banks continuing to lend. In this equilibrium, which arguably is the most interesting one, the default probability has been shown to be decreasing in the proportion of honest borrowers, the number of competent banks and the quality of the screening technology.

Economic growth—measured by an increase in the proportion of borrowers who have access to an investment opportunity—can reduce the default rate in both the Low and High Default equilibria. In the latter, however, the reduction of the default rate will be greater, the larger the increase in honest borrowers gaining access to the new investment opportunity.

We have presented empirical evidence from a large number of African banks that is consistent with our theoretical predictions. Specifically, we have shown that growth is inversely related to the rate of loan defaults, as does the rule of law and control of corruption. Bank age, which we consider a good proxy for a banks screening ability, rapidly reduces the default rate, but only once a bank is no longer a ‘baby’. Government banks in corrupt environments have been found to experience higher loan defaults than similar age private banks. On the other hand, when control of corruption is above the world average, government ownership is found to reduce the default rate.

Our results have straightforward policy implications. Improved screening of borrowers, which calls for the development of credit bureaus and better information sharing, seems critical to reduce adverse selection. Better governance, especially better contract enforcement and control of corruption, seems equally important in terms of deterring moral hazard by opportunistic borrowers. Finally, although it does appear that government banks have better information capital and can, in principle, reduce loan defaults, government ownership appears to make matters worse in countries in which control of corruption is below the world norm.

References

Andrianova, S., B. Baltagi, P. Demetriades, and D. Fielding (2011) 'Why do African banks lend so little?' University of Leicester Discussion Paper in Economics, # 11/19

Demetriades, P., and D. Fielding (2010) 'Information, institutions and banking sector development in West Africa.' *Economic Inquiry.* DOI: 10.1111/j.1465-7295.2011.00376.x

Goodhart, C., and D. Tsomokos (2011) 'The role of default in macro-economics.' Mimeo

Gourieroux, C.S., A. Monfort, and A. Trognon (1984) 'Pseudo maximum likelihood methods: Theory.' *Econometrica* 52, 681–700

Honohan, P., and T. Beck (2007) *Making Finance Work for Africa* (Washington, DC: The World Bank)

Hotelling, H. (1929) 'Stability in competition.' *Economic Journal* 39, 41–57

McCullagh, P., and J.A. Nelder (1989) *Generalised Linear Models* (London: Chapman & Hall)

Papke, L. E., and J. M. Wooldridge (1996) 'Econometric methods for fractional response variables with an application to 401(k) plan participation rates.' *Journal of Applied Econometrics* 11, 619–32

Rousseau, P.L. (1998) 'The permanent effects of innovation on financial depth: Theory and US historical evidence from unobservable components models.' *Journal of Monetary Economics* 42, 387–425

Shubik, M. (1973) 'Commodity money, oligopoly, credit and bankruptcy in a general equilibrium model.' *Western Economic Journal* 10, 24–38

Stiglitz, J., and A. Weiss (1981) 'Credit rationing in markets with imperfect information.' *The American Economic Review* 71, 393–410

Appendix A: Theory proofs

The proof of the proposition sets out the payoffs of all players and then establishes the conditions which deliver the stated pooling equilibria.

The expected payoff of a borrower of each type from applying for a loan to bank $i = \{A, B\}$ is as follows:

$$\begin{aligned}
U_i^\alpha &= [\kappa \cdot (\xi + (1-\xi)p_\kappa) + (1-\kappa)p_{1-\kappa}][R - r_i] - tx_i^\alpha \\
U_i^\beta &= [\kappa \cdot (\xi(1-\sigma) + (1-\xi)p_\kappa) + (1-\kappa)p_{1-\kappa}] - tx_i^\beta \\
U_i^\gamma &= [\kappa \cdot (\xi(1-\sigma) + (1-\xi)p_\kappa) + (1-\kappa)p_{1-\kappa}][(1+R)(1-\lambda(1-q)) - q(1+r_i)] - tx_i^\gamma
\end{aligned}$$

where $x_i^{\{.\}}$ stands for the distance between the borrower of type $\{.\}$ and bank i, while $x_B^{\{.\}} = 1 - x_A^{\{.\}}$. The payoff to a bank of the given type is written as:

$$\begin{aligned}
V_i^\kappa &= D_i\Big[\xi\{(1+r_i)[\alpha + \gamma(1-\sigma)(q + \lambda(1-q))] + (1+r_0)\sigma(\gamma+\beta)\} + \\
&\quad + (1-\xi)\{p_\kappa(1+r_i)[\alpha + \gamma(q + \lambda(1-q))] + (1-p_\kappa)(1+r_0)\}\Big] \\
V_i^{1-\kappa} &= D_i\Big[p_{1-\kappa}(1+r_i)[\alpha + \gamma(q + \lambda(1-q))] + (1-p_{1-\kappa})(1+r_0)\Big]
\end{aligned}$$

where D_i is the demand for bank i loan contracts.

LDE is defined as an equilibrium with $q^* = 1$, $\xi^* = 1$ and $p_{1-\kappa}^* = 1$. For $q^* = 1$, we check that γ-borrower will not want to deviate by $q = 0$ when $\xi^* = 1$ and $p_{1-\kappa}^* = 1$:

$$\begin{aligned}
U_i^\gamma(q = 1|\xi = 1, p_{1-\kappa} = 1) &\geq U_i^\gamma(q = 0|\xi = 1, p_{1-\kappa} = 1) \\
(1-\kappa\sigma)(R - r_i) - tx_i^\gamma &\geq (1-\kappa\sigma)(1-\lambda)(1+R) - tx_i^\gamma \\
\lambda &\geq \frac{1+r_i}{1+R} \qquad (1)
\end{aligned}$$

The competent bank will choose $\xi^* = 1$ when $V_i^\kappa(\xi = 1|q^* = 1, p_{1-\kappa}^*) \geq V_i^\kappa(\xi = 0|q^* = 1, p_{1-\kappa}^*)$, or equivalently when

$$\beta \geq \gamma(r_i - r_0)/(1+r_0). \qquad (2)$$

In order to find the equilibrium value of r_i, write the total demand for bank i loan contracts as

$$D_i = \alpha D_i^\alpha + \beta D_i^\beta + \gamma D_i^\gamma \qquad (3)$$

i.e. it is the sum of total demands per type of borrower. These latter ones are determined by the marginal borrower of each type. Each type marginal borrower is indifferent between going to bank A or bank B for a loan. For the honest marginal borrower this gives

$$x_A^{\alpha} = \frac{1}{2} - \frac{r_A - r_B}{2t} \tag{4}$$

Similarly, the marginal opportunistic borrower is given by

$$x_A^{\gamma} = \frac{1}{2} - \frac{r_A - r_B}{2t}(1 - \kappa\sigma) \tag{5}$$

If the dishonest borrower located exactly in the middle of the interval between the two banks has a non-negative payoff, then every dishonest borrower will apply to the nearest bank. This translates into

$$x_A^{\beta} = 1/2 \text{ when } \kappa\sigma \geq 1 - t/2. \tag{6}$$

Collecting the terms and making the required assumptions, we have

$$D_A = \frac{1}{2} - \frac{1}{2t}[\alpha + \gamma(1 - \kappa\sigma)](r_A - r_B) \tag{7}$$

Substituting this into competent bank's payoff and solving the first order condition for a symmetric solution ($r_A = r_B$), it can be checked that

$$1 + r_A^* = 1 + r_B^* = \frac{t}{\alpha + \gamma(1 - \kappa\sigma)} - \frac{\sigma(\gamma + \beta)(1 + r_0)}{\alpha + \gamma(1 - \sigma)} \tag{8}$$

To ensure that all opportunistic borrowers apply for a loan (i.e. that the marginal opportunistic borrower is located in the middle of the interval), it is sufficient to assume that $t \leq \sigma(1 - \alpha)(1 + r_0)$. Note that when the participation constraint of opportunistic marginal borrower is staisfied, so will be the PC of the honest borrower (because the expected payoff of an honest borrower in LDE is higher than that of an opportunistic borrower located at the same point). The stricter of the two conditions on t will ensure that borrowers of every type apply.

To solve for HDE with $q^* = 0$, $\xi = 1$ and $p_{1-\kappa} = 1$, repeat the steps of the solution for LDE. Opportunistic borrowers chose $q^* = 0$ when the reverse of (1) holds.

The competent type of bank prefers to screen all its loan applications if (2), as before. Additionally, in this case, given that $q^* = 0$, the competent bank prefers screening and lending to those with untainted record over not screening and not lending to any borrower: $V_i^\kappa(\xi^* = 1|q^* = 0) \geq V_i^\kappa(\xi^* = 0, p_\kappa = 0|q^* = 0)$, which obtains when

$$\lambda \geq \frac{(1-\sigma(1-\alpha))(1+r_0)}{\gamma(1-\sigma)(1+r_A)} - \frac{\alpha}{\gamma(1-\sigma)} \tag{9}$$

Since opportunistic borrowers do not repay their loans in HDE, their expected payoff no longer depends on r_i and therefore the marginal borrowers of each type in HDE are given by (4), (6) and

$$x_A^\gamma = 1/2 \text{ when } (1-\kappa\sigma)(R-r_A) \geq t/2 \tag{10}$$

Solving for r_A from the first order condition of the expected payoff maximisation of the competent bank and assuming a symmetric solution, the equilibrium rate in HDE is

$$1 + r_A^* = 1 + r_B^* = \frac{2t}{\alpha} - \frac{\sigma(1-\alpha)(1+r_0)}{\alpha + \gamma\lambda(1-\sigma)} \tag{11}$$

To complete the proposition, NLE obtains when the competent bank finds it more profitable to invest the loanable funds into the safe asset rather than to make loans: $V_i^\kappa(\xi^* = 1|q^* = 0) < V_i^\kappa(\xi^* = 0, p_\kappa = 0|q^* = 0)$, which is the reverse of (9).

Appendix B: Empirics

TABLE 1: *Summary Statistics (110 banks from 29 countries)*

Variable name	**Number of Observations**	**Mean**	**Standard Deviation**	**Minimum**	**Maximum**
Default rate	519	0.1086	0.1311	0.0001	0.8087
Bank age	519	36.7861	32.7211	3	147
Government share	519	7.8566	20.0841	0	100
Logarithm of assets	519	6.5584	2.8756	–1.5539	20.2198
Control of corruption	519	–0.4659	0.6464	–1.5464	1.0708
Regulatory quality	519	–0.4203	0.7153	–2.3694	0.9536
Rule of law	519	–0.5330	0.6654	–1.7216	0.8956
Growth rate	519	0.0220	0.0484	–0.1644	0.2541

TABLE 2: *Correlation Matrix*

	Default rate	Bank age	Government share	Logarithm of assets	Control of corruption	Regulatory quality	Rule of law
Bank age	–0.0452						
Government share	0.2134	–0.0249					
Logarithm of assets	–0.1142	0.3149	–0.0089				
Control of corruption	–0.1652	0.0222	0.0662	–0.0216			
Regulatory quality	–0.0353	0.0072	0.0585	–0.2151	0.8195		
Rule of law	–0.1001	0.0074	0.1217	–0.0691	0.9266	0.8546	
Growth rate	–0.0621	–0.1701	0.0867	–0.2431	0.2497	0.4178	0.3315

TABLE 3: *OLS and Random Effects (RE) regressions of the default rate*

	All Countries		Sub-Saharan Africa	
	Method of Estimation		Method of Estimation	
Regressor	**OLS**	**RE**	**OLS**	**RE**
Bank age	0.1108**	0.2292**	0.0648	0.1849**
	(0.0484)	(0.1081)	(0.0459)	(0.0910)
Bank age squared	–0.0991***	–0.1771**	–0.0453	–0.1278**
	(0.0333)	(0.0767)	(0.0307)	(0.0640)
Logarithm of assets	–0.0025*	–0.0049**	–0.0024*	–0.0033
	(0.0015)	(0.0025)	(0.0015)	(0.0022)
Government share	0.0455*	0.0800	0.0993***	0.1234***
	(0.0271)	(0.0510)	(0.0299)	(0.0462)
Rule of law	–0.0480**	–0.0642***	–0.0340*	–0.0775**
	(0.0233)	(0.0256)	(0.0202)	(0.0331)
Control of corruption	–0.0795***	–0.0406**	–0.0999***	–0.0579**
	(0.0190)	(0.0187)	(0.0185)	(0.0247)
Regulatory quality	0.0824***	0.0606***	0.0815***	0.0738***
	(0.0185)	(0.0228)	(0.0182)	(0.0253)
Growth rate	–0.4096***	–0.7156	–0.4585***	–0.0578
	(0.1026)	(0.0845)	(0.1089)	(0.0853)
North Africa	0.1149***	0.1055***		
	(0.0171)	(0.0412)		
Constant	0.0954***	0.0954***	0.0442**	0.0077
	(0.0196)	(0.0196)	(0.0194)	(0.0267)
Interaction terms				
Government share x	–0.2635***	–0.1355	–0.0379***	–0.2572***
control of corruption	(0.0721)	(0.0859)	(0.0833)	(0.0969)
Rule of law x bank age	0.0412***	0.0004***	0.0005***	0.0005***
	(0.015)	(0.0001)	(0.0001)	(0.0002)
Summary Statistics and Diagnostics				
No. of obs	519	519	442	442
No. of banks	110	110	95	95
Joint significance				
of time dummies	3.43	33.02	3.40	29.41
[p-value]	[0.001]	[0.000]	[0.001]	[0.000]
Rho (Fraction of				
variance due to RE		0.7083		0.6103
Breusch-Pagan LM test for RE		403.75		328.09
[p-value]		[0.000]		[0.000]

Note: All regressions include a full set of time dummies

Table 4: *GLM regressions of the default rate*

	All Countries		Sub-Saharan Africa	
	Bernoulli Binary Family		Bernoulli Binary Family	
	Link Function		Link Function	
Regressor	**Logistic**	**Log Log**	**Logistic**	**Log Log**
Bank age	1.995***	1.943***	1.589**	1.562**
	(0.632)	(0.599)	(0.710)	(0.678)
Bank age squared	–1.804***	–1.755***	–1.006**	–1.001**
	(0.466)	(0.445)	(0.469)	(0.450)
Logarithm of assets	–0.039*	–0.037*	–0.029	–0.026
	(0.020)	(0.020)	(0.020)	(0.019)
Government share	–0.143	–0.272	0.551	0.555
	(0.507)	(0.505)	(0.471)	(0.438)
Rule of law	–0.651***	–0.651***	–0.654**	–0.715***
	(0.254)	(0.239)	(0.294)	(0.290)
Control of corruption	–0.917***	–0.833***	–1.283***	–1.145***
	(0.192)	(0.178)	(0.234)	(0.218)
Regulatory quality	0.885***	0.844***	0.945***	0.910***
	(0.185)	(0.178)	(0.195)	(0.190)
Growth rate	–3.638***	–3.353***	–3.926***	–3.321***
	(1.124)	(1.062)	(1.221)	(1.112)
North Africa	1.380***	1.283***		
	(0.161)	(0.150)		
Constant	–3.426***	–3.414***	–3.886***	–3.933***
	(0.275)	(0.264)	(0.365)	(0.357)
Interaction terms				
Government share x	-2.634***	–2.536***	–3.317***	–2.818***
control of corruption	(0.945)	(0.953)	(0.849)	(0.753)
Rule of law x bank age	0.624***	0.595***	1.077***	0.996***
	(0.235)	(0.226)	(0.284)	(0.278)
Summary Statistics and Diagnostics				
No. of obs	519	519	442	442
No. of banks	110	110	95	95
Joint significance				
of time dummies	25.14	24.01	29.09	29.30
[p-value]	[0.001]	[0.001]	[0.000]	[0.000]
Log pseudolikelihood	–125.27	–125.34	–97.60	–97.61
AIC	0.556	0.556	0.523	0.523
BIC	–3077.9	–3077.8	–2546.7	–2546.7

Note: All regressions include a full set of time dummies.