



Improving Credit Risk Analysis with Psychometrics in Peru

Bailey Klinger
Asim Ijaz Khwaja
Joseph LaMonte

**Inter-American
Development Bank**

Capital Markets and
Financial Institutions
Division (CMF),
Institutions for
Development (IFD)

TECHNICAL NOTE

No. IDB-TN-587

October 2013

Improving Credit Risk Analysis with Psychometrics in Peru

Bailey Klinger
Asim Ijaz Khwaja
Joseph LaMonte



Inter-American Development Bank

2013

Cataloging-in-Publication data provided by the
Inter-American Development Bank
Felipe Herrera Library

Klinger, Bailey.

Improving credit risk analysis with psychometrics in Peru / Bailey Klinger, Asim Ijaz Khwaja, Joseph LaMonte.

p. cm. (IDB Technical Note; 587)

Includes bibliographical references.

1. Small business—Finance—Peru. 2. Credit analysis—Peru. I. Khwaja, Asim Ijaz. II. LaMonte, Joseph. III. Inter-American Development Bank. Capital Markets and Financial Institutions Division. IV. Title. V. Series.

IDB-TN-587

<http://www.iadb.org>

The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent.

The unauthorized commercial use of Bank documents is prohibited and may be punishable under the Bank's policies and/or applicable laws.

Copyright © 2013 Inter-American Development Bank. All rights reserved; may be freely reproduced for any non-commercial purpose.

Abstract

Access to finance remains a challenge for some micro, small, and medium-sized enterprises (MSMEs) in Peru, particularly informal enterprises with no borrowing history in the formal financial system. Lenders lack the tools to reach these borrowers with sufficient scale and control over risk due in part to the shortcomings of current screening technologies. For this paper, the authors carried out a pilot test of an innovative psychometric tool aimed at evaluating credit risk for business owners seeking a loan from Financiera Confianza, the fourth largest Empresa Financiera in Peru. Applicant responses were compared to self-reported sales, subsequent loan repayment performance, and credit bureau data to determine if psychometric-based credit scoring models could reduce the constraints on MSME finance. The authors created a scorecard based on that information using data from other countries and evaluated its effectiveness on this sample. It achieved a Gini coefficient of between 20 and 40 percent. Those MSMEs rejected by a psychometrically enhanced application scorecard with this Gini coefficient have a probability of defaulting that is up to four times greater than those accepted by the scorecard. Along with other policies to reduce information asymmetry in MSME lending, such a tool could help relieve constraints on MSME finance in Peru.

Keywords: MSME, credit, risk, scoring, Peru, bank, finance, psychometrics

JEL Codes: G21, G32, H32

Table of Contents

Figures and Tables.....	3
Introduction.....	4
Access to Finance in Peru.....	5
Credit Risk Systems in Peru	12
Barriers in Risk Assessment for MSMEs	15
Pilot Trial Implementation.....	18
Results	22
Conclusions and Further Directions	38
References.....	40
Appendix: Graphing and Interpreting AUC and Gini Coefficient.....	42

Figures and Tables

Figure 1. Domestic Credit as Percent of GDP in Peru, 2000–12.....	5
Figure 2. Financial Depth as Percentage of GDP, 2010	6
Figure 3. Percent of Firms that Identify Access to Finance as a Major Constraint, 2010 ..	7
Figure 4. Firms with a Bank Loan or Credit Line by Years in Operation, 2010	8
Figure 5. Firms by Size Using Bank Loans to Finance Investment (percent), and Working Capital Investments Financed with Bank Loans by Firm Size (percent), 2010	9
Figure 6. Percent of Formal and Informal Firms Identifying Access to Finance as a Major Constraint, 2010	10
Figure 7. Digit Span Recall Results by Country	24
Figure 8. Digit Span Recall Results and Monthly Sales by Country	25
Figure 9. Percentage of the Sample that Agrees with Each Statement by Country	26
Figure 10. Percentage of the Sample that Agrees with Each Statement by Country vs. Business Sales	28
Figure 11. Correlation between Specific Personality Items, Intelligence, and Business Revenues by Country	30
Figure 12. Correlation of Demographics with Business Revenues	31
Figure 13. Model Built on Financiera Confianza Data	34
Figure 14. ROC Curve - African Model Applied to African Data	35
Figure 15. African Model on SBS 60Days+ Arrears for Financiera Confianza	36
Figure 16. African Model on SBS 60Days+ Arrears for Financiera Confianza vs. Volume by Score Cut-off (by decile)	37
Figure 17. Building an ROC Curve	42
Figure 18. ROC Curve of a Perfect Model	43
Figure 19. ROC Curve of a Random Model	44
Table 1. Cross-Country Summary Statistics for Comparator Countries	23
Table 2. AUC & Gini of African Model Applied to Financiera Confianza	35

Introduction

By many measures, the Peruvian financial system is highly effective. The regulatory regime is internationally lauded, and the industry is quite open, with high levels of competition, which has lowered interest rates. As a result, few formal firms characterize the cost of or access to finance as a constraint on their business compared to other countries in the Latin American region. However, challenges remain. As a whole, the financial sector is shallow, with credit to the private sector as a percentage of GDP less than half that of the rest of the countries in the region. Geographically, many areas are underserved. Most significantly, small informal enterprises are typically denied access to finance unless they have collateral and/or a borrowing history, and they face very high interest rates.

As the informal sector is a significant proportion of the Peruvian economy, this represents a major development challenge for the country, as well as an attractive business opportunity for the increasingly crowded and competitive financial services industry. But even the large and innovative microfinance sector in Peru is unable to effectively reach these smaller, largely informal and unbanked applicants due to a combination of high operational expenses and little available information for risk analysis.

This paper proposes a new tool to overcome these barriers and help capture this opportunity: a psychometrically enhanced credit application scorecard. This tool was tested on a sample of roughly 275 enterprises borrowing from a leading Peruvian microfinance institution from May to December 2012. The goal was to determine if, and to what degree, this tool could discriminate an enterprise's level of risk to the lender. The results suggest a powerful contribution to risk analysis that is highly scalable and can be implemented even on unbanked business owners with no credit history or collateral.

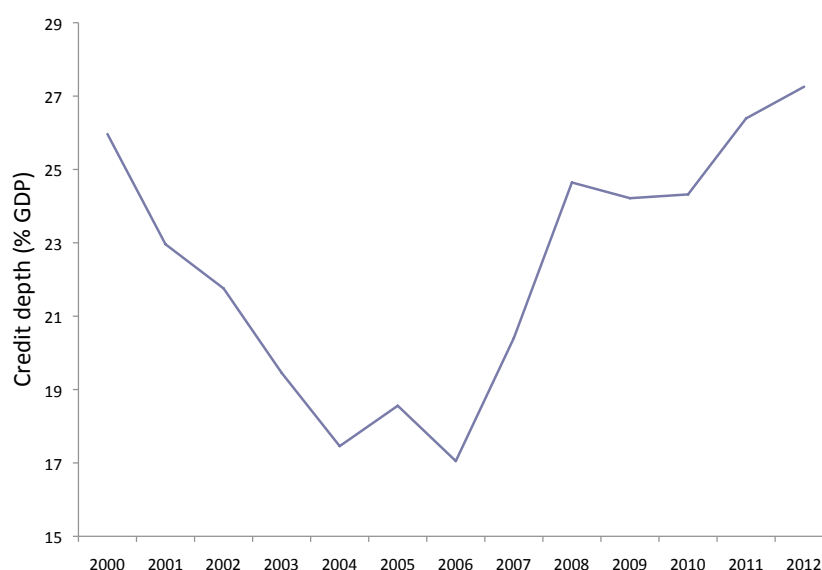
The next section of this paper describes the characteristics of Peru's financial system, with particular attention to the situation facing micro, small, and medium-sized enterprises (MSMEs). The third section analyses existing credit risk systems for MSME applicants, particularly for Financiera Confianza, a leading Empresa Financiera that participated in our pilot test of the psychometric tool. The potential of this tool to solve challenges facing financial institutions like Financiera Confianza is described in the section after that. We then review the methodology of the pilot trial and present detailed results, including cross-country comparisons

with other pilot studies and an evaluation of the tool's predictive power on Financiera Confianza's clients. The final section concludes.

Access to Finance in Peru

Since 2001, the Peruvian economy has experienced a period of rapid and sustained economic growth marked by increases in internal demand and in international prices for minerals and nontraditional exports. Between 2001 and 2010, GDP grew at an average rate of 6.30 percent (INEI, 2011). Yet, in the face of such robust economic growth, domestic credit as a percentage of GDP, while rebounding from a low of 17.05 percent in 2006, was only 27.26 percent in 2012 (Figure 1).

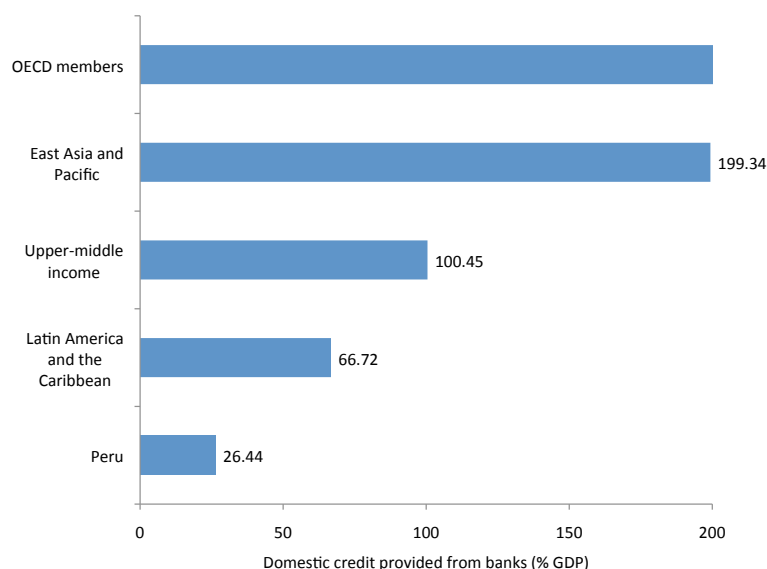
Figure 1. Domestic Credit as Percent of GDP in Peru, 2000–12



Source: World Bank (2012a).

Using this measure, Peru's financial depth is less than half that of the rest of the countries in the Latin American region (65 percent), four times less than that of economies with a similar income per capita (100.45 percent), and eight times less than that of OECD member countries (199.34 percent) and East Asia and Pacific (202.20 percent) (Figure 2). In turn, bank deposits from the private sector stand at 19.6 percent of GDP. Nearly 40 percent of deposits are denominated in dollars, a number that has declined from a high of 70 percent over the last decade as Peru's currency has appreciated in value against the dollar (Oxford Business Group, 2012).

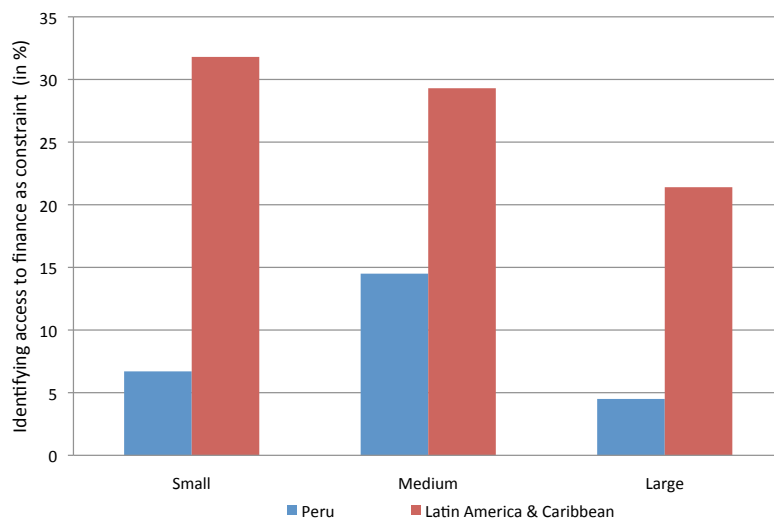
Figure 2. Financial Depth as Percentage of GDP, 2010



Source: Authors' calculations based on World Bank (2010a; 2010b).

However, contradicting Peru's low domestic credit as a percent of GDP, access to and use of finance is perceived to be relatively high, across all business sizes, compared to Latin America and the Caribbean (LAC) as a whole. Only 8.5 percent of Peruvian companies formally registered with the taxation authority identify finance as a major constraint on investment, and the measure is only slightly higher among medium (14.5 percent) and small companies (6.7 percent) than among large ones (4.5 percent) (World Bank, 2010a). In contrast, across LAC, 30.8 percent of formally registered companies identify access to finance as a major constraint to investment, with small (31.8 percent), medium (29.3 percent), and large (21.4 percent) enterprises all having considerably more difficulty accessing finance than similarly sized Peruvian firms (World Bank, 2010a) (Figure 3).

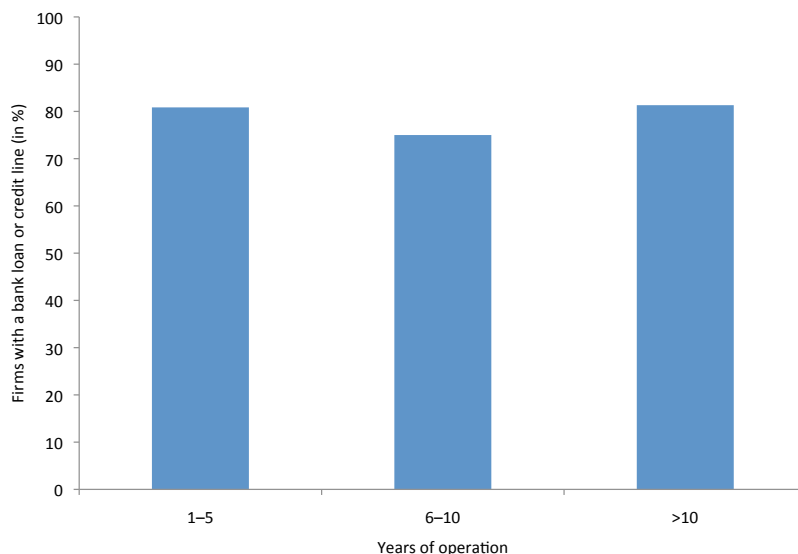
Figure 3. Percent of Firms that Identify Access to Finance as a Major Constraint, 2010



Source: World Bank (2010a).

In Peru, distribution of credit to the private corporate sector is slightly uneven on a territorial level. At the department capital level, businesses in Huancayo (49.8 percent), Trujillo (49.4 percent), Arequipa (43.4 percent), and Piura (40.7 percent) have greater financial access than those in Lima-Callao (35.9 percent), Chiclayo (35.1 percent), and Iquitos (29.9 percent) (INEI, 2011). Additionally, among formal firms, access to finance is largely unrelated to the age of the enterprise in Peru. As seen in Figure 4, 81 percent of companies with between one and five years of operations or 10 or more years of operations have a bank loan or credit line, and 75 percent of enterprises with between five and 10 years of operations are able to obtain bank loans.

Figure 4. Firms with a Bank Loan or Credit Line by Years in Operation, 2010



Source: Authors' calculations based on World Bank (2010a).

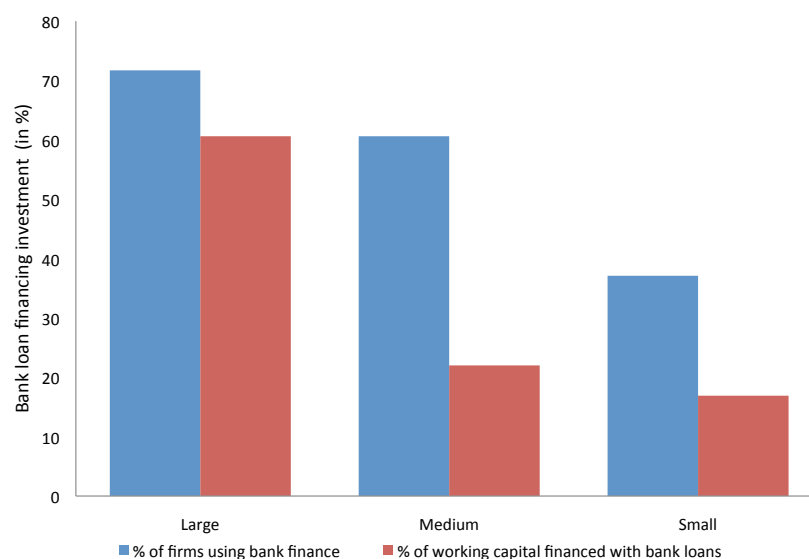
This could be considered a contradiction: financial depth in Peru is extremely low, yet when formal enterprises are asked about access to finance, they do not perceive it as a problem. This could be explained by (a) alternative sources of finance for Peruvian firms and (b) differences in access to finance between the formal and informal economy.

Formal Peruvian firms finance 41.8 percent of their working capital investments with internal funds and retained earnings (authors' calculations based on World Bank, 2010a). Among formal enterprises, 51.8 percent of small and 29.7 percent of medium enterprises' investments are financed internally, while large firms rely on internal finance for 37.0 percent of their investments (World Bank, 2010a). Across firm sizes, formal enterprises only finance 29.4 percent of their working capital with bank loans, with the remainder coming from suppliers (23.6 percent) and other sources, including microfinance institutions (MFIs) (5.2 percent) (World Bank, 2010a).

The probability of taking bank loans and the proportion of working capital financed through bank loans is directly related to a firm's size. Of the 29.4 percent of working capital that formal firms finance through bank loans, 71.7 percent of large-sized firms use banks to finance 37.1 percent of their working capital, 60.6 percent of medium-sized enterprises use banks to

finance 22.0 percent of their working capital, and 37.1 percent of small enterprises use banks to finance 16.9 percent of their working capital (Figure 5) (World Bank, 2010a).

Figure 5. Firms by Size Using Bank Loans to Finance Investment (percent), and Working Capital Investments Financed with Bank Loans by Firm Size (percent), 2010



Source: World Bank (2010a).

Formal small and medium-sized enterprises (SMEs) especially benefit from the increasingly competitive nature of the Peruvian financial system, which has grown substantially since 1999. At the close of 2011, more than 40 institutions catered to the financial needs of small businesses. The resulting competition pushed average interest rates to small businesses down from 73 percent in 1998 to 32 percent in 2009 (Quispe, León, and Contreras, 2012).

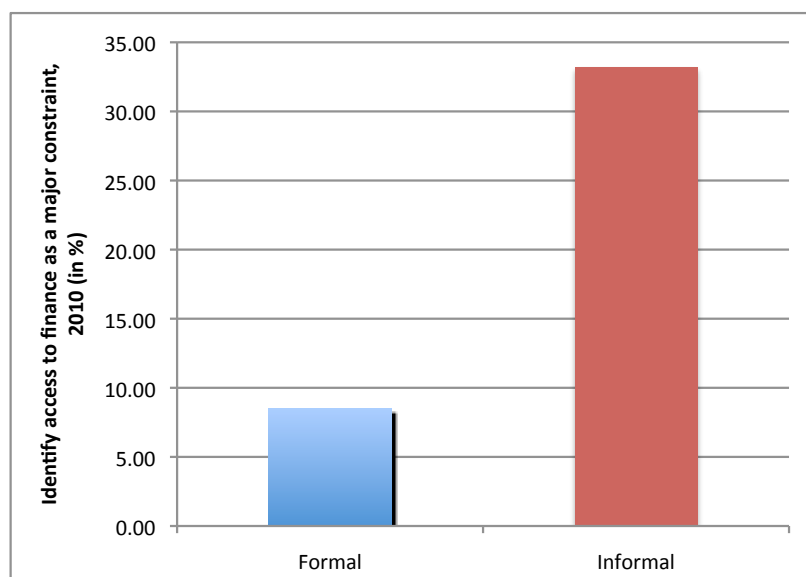
Despite this high level of competition, informal businesses in Peru have much more difficulty accessing finance than formal businesses, both in terms of general access and available product offerings. Peru's informal sector, while shrinking, remains large in both absolute and relative terms. According to Peru's national household survey (La Encuesta Nacional de Hogares), in 2008, there were 3,383,325 micro and small businesses, of which 2,264,071, or 66.9 percent, were informal (CODEMYPE, 2011). The World Bank Enterprise Survey for 2010 showed that 68.6 percent of firms in Peru compete against unregistered or informal firms and

37.3 percent of firms identify practices of competitors in the informal sector as a major constraint on business (World Bank, 2010a).

Informality creates impediments to accessing finance, as lack of information about business projects, sales, and growth potential, and the lack of transparency in balance sheets make it difficult for banks to assess the risk of default of a given business (Ferraro, 2011). To compensate, lending institutions often enforce the delivery of collateral, which many small businesses do not have, and/or charge high interest rates to cover the risk (Ferraro, 2011). A study by San Ignacio de Loyala University found that the average interest rate charged by banks to small businesses was 32 percent, while loans from nonregulated institutions, such as nongovernmental organizations (NGOs), savings and loan cooperatives, and informal lenders—the traditional loan sources for many informal businesses—range between 69 and 588 percent annually (Quispe, León, and Contreras, 2012).

Such factors cause informal enterprises to have greater difficulty accessing finance than formal businesses, both in terms of access generally and in terms of product offerings available. Whereas 8.5 percent of formally registered Peruvian firms identify access to finance as a major constraint to investment, 33.2 percent of informal firms identify it as a severe obstacle to current operations (Figure 6) (World Bank, 2010a and 2010b).

Figure 6. Percent of Formal and Informal Firms Identifying Access to Finance as a Major Constraint, 2010



Source: Authors' calculations based on World Bank (2010a; 2010b).

Among informal firms, 86.3 percent finance their working capital with internal funds, yet 43.5 percent of those firms are dissuaded from seeking a loan for a variety of reasons (authors' calculations, based on World Bank, 2010b). In one recent nationwide survey, informal businesses rated high interest rates (43.6 percent), complex application procedures (34.6 percent), and lack of collateral (16.0 percent) as their primary reasons for not seeking a loan (author's calculations based on World Bank, 2010b). In a separate survey, micro and small businesses (the majority of which are informal) rated lack of collateral (32.7 percent), lack of adequate documentation (24.6 percent), and inability to demonstrate company revenues (20.0 percent) as their main impediments to accessing finance (INEI, 2011). Of note, roughly 14.5 percent of Peruvian firms registered their companies specifically to obtain better access to finance (authors' calculation based on World Bank, 2010a).

The cumulative impact of such market factors is that only 30.3 percent of micro and small businesses receiving finance in Peru are satisfied with the services they receive, and such businesses rank interest rates (47.5 percent) and repayment terms (22.5 percent) as the most important characteristics for evaluating a loan offer (INEI, 2011).

In summary, financial depth in Peru is significantly lower than for its Latin American peers, with private sector credit as a percentage of GDP less than half of that found across the region. Relative openness of the financial sector combined with good macroeconomic performance and an effective and efficient regulatory system (discussed in the following section) have led to increased supply of credit and reductions in interest rates. As a result, most formal firms surveyed do not see access to finance as a barrier to expansion. However, it appears that this high level of access has not reached many geographic regions of the country and does not extend to informal firms, which represent a significant proportion of employment and economic activity in Peru. Informal enterprises face restricted access to more expensive loans. As a result, many are forced to self-finance their operations and consider their lack of access to finance as a severe obstacle to their business's current operations.

Credit Risk Systems in Peru

This section presents an overview of regulation, oversight, and functioning of credit risk systems for commercial loans in Peru. In addition, a brief description of Financiera Confianza's recent history and performance is given, along with an abridged depiction of its credit risk assessments for loans to private companies.

Credit Risk for Commercial Loans

The regulation and supervision of Peru's financial system is primarily the responsibility of the functionally autonomous superintendent of banking, pensions, and insurance (Superintendencia de Banca, Seguros y AFP, or SBS), the entity through which Peru has earned a reputation as a regional leader in terms of sound banking and economic regulation. In 2005, the SBS was rated 96.6 out of 100 by a combined World Bank-International Monetary Fund mission for the quality of its general financial regulation and supervision. Also, the Economist Intelligence Unit (EIU) gave Peru the highest possible score for its regulatory and examination capacity (EIU, 2009).

Among its responsibilities established in the *General Law of the Financial and Insurance Systems*, the SBS protects the public interest by ensuring the stability and solvency of the firms in Peru's financial system. Financial institutions must comply with the SBS's regulatory framework, which requires such institutions to file monthly reports that comply with the agency's auditing and reporting requirements.

The SBS requires financial institutions to formally assess the financial situation of potential clients before issuing loans. Though not explicitly stated, the goal of such assessments is to prevent clients from becoming overly indebted. Thus, after collecting monthly reports from Peru's financial institutions, the SBS compiles individual-specific reports that show a potential client's level of indebtedness. Specifically, the SBS reports (a) show the number of active loans from SBS-regulated institutions that an individual has had during each month of the previous year; (b) classify the percentage of such debt that falls into five repayment categories, each of which occupies a range of days late the individual paid existing loans during each month of the previous year; and (c) indicate the amount of debt falling into different classifications, including working capital debt, consumer debt, mortgage debt, and refinanced debt.

Recognizing the diverse nature of lending, the SBS allows each financial institution to establish its own client background information requirements and risk analyses when making a

lending decision. Peruvian banks thus employ different credit risk methods, which can be automatic, semiautomatic, or manual. Evaluating commercial loans usually involves considering liquidity, financial structure, payment behavior, application fraud checks, governance and management standards, information technology systems, outlook for the borrower's main activity sector, the firm's legal status, and the existence of refinancing procedures or write-offs. Some banking institutions use econometric techniques to estimate credit scores and default probabilities. Other financial institutions use risk matrices based on companies' financial information, while still others rely on analysis by credit officials based on simple rules of thumb to determine whether to approve a loan.

In obtaining system-wide client reports, the SBS consolidates and shares the information with all financial institutions through its credit bureau. Such a system ensures that all financial institutions are informed about a potential borrower's state of indebtedness.

Financiera Confianza

Financiera Confianza is a privately operated microfinance operation located in several provinces in Peru. The origin of the institution dates to 1992, when it was created through an agreement between the Inter-American Development Bank's (IDB) Small Projects Program and the NGO for education services promotion and rural support (Servicios Educativos Promocion y Apoyo Rural, or SEPAR). The organizations teamed up to create a microcredit program that sought to improve wages of low-income women living in Junin and Huancavelica, two of Peru's most impoverished regions.

In September 1997, the original operation grew into an Entity for the Development of Small and Micro Enterprises (Entidad de Desarrollo para la Pequeña y Microempresa, or EDPYME), a credit-only institution designed to formalize loan-granting NGOs, and began operations in Peru's central region. Then, in October 2009, Financiera Confianza obtained the SBS's approval to transform from an EDPYME to Empresa Financiera (EF), which allows Financiera Confianza to receive deposits from the public while issuing loans in compliance with the SBS's regulations. Currently, Financiera Confianza has 40 branches scattered across seven departments that adjoin greater Lima and Peru's central region.

Financiera Confianza has successfully leveraged its experience working with MSMEs. Since initiating operations as a financial company, the institution has grown into Peru's fourth largest EF as measured by total lending, deposits, and equity.

Additionally, Financiera Confianza's proportion of nonperforming loans to total lending is stellar (1.64 percent), well below the sector average for all financial companies (5.17 percent). Presently, Financiera Confianza directs more than 75 percent of its lending to MSMEs and the rest to personal loans and mortgages. Since December 2009, its lending to the private sector has increased by 167 percent, and the number of loans issued to MSMEs has increased by 173 percent.¹

Financiera Confianza's Credit Risk Analysis for Commercial Loans

Following SBS regulations, Financiera Confianza has created background information requirements that vary based on the loan product that applicant requests. Overall, the institution's analysis considers concrete, quantitative variables along with qualitative factors. The quantitative analysis includes a series of financial variables, such as demonstrated or developed cash flows, profits, current debt in relation to revenues, and credit bureau analysis, which is the applicant's repayment history, the number of entities in which the applicant has a loan, and the pattern of debt increase or decrease over the year prior to applying for a loan. Additionally, Financiera Confianza considers other concrete variables aimed at demonstrating the potential client's stable presence, such as water or electricity payment receipts and rental or mortgage contracts. Overall, such factors are accorded roughly a 70 percent weight when determining the risk associated with each loan application. Financiera Confianza also requires a guarantee for most loans above 2,000 Peruvian Nuevos Soles (approximately US\$750), which in many cases takes the form of a stake in the merchandise purchased through the loan.

In addition to quantitative factors, Financiera Confianza considers several qualitative variables in assessing a loan applicant's credit risk. Standard protocol requires that loan analysts visit the applicant's place of business and converse with neighbors to ascertain his or her reputation in the community. Additionally, Financiera Confianza assesses the safety of the applicant's neighborhood and proximity in relation to a Financiera Confianza office, key factors related to collecting payments from clients. Such factors are accorded roughly a 30 percent weight when determining the risk associated with each loan application.

¹ Data in this paragraph are based on SBS data from March 2013.

It is noteworthy that the factors are not specifically weighted, but rather provide general guidelines for credit approval. The final decision comes from Financiera Confianza's office-specific loan approval committee, which has discretion to balance the factors considered as it sees fit.

Barriers in Risk Assessment for MSMEs

Credit risk analyses by lenders such as Financiera Confianza reflect the typical approach and offer numerous advantages, but continue to face some challenges. This section summarizes these challenges—faced in the region and around the world—which result in higher costs and limited access for borrowers and squeezed margins and slow growth for financial institutions, and shows how leveraging psychometric application content could help relieve some of them.

Shortcomings of the Current Approach

Financiera Confianza's approach has clearly produced effective results in terms of maintaining a portfolio with a proportion of nonperforming loans to total lending well below the sector average. However, this approach suffers from several weaknesses.

First, under Financiera Confianza's approach, it is difficult to identify and assess potential clients who have never previously taken a loan from a financial institution regulated by the SBS. The SBS data on which the institution bases many aspects of its individual analysis only exist if that individual has previously taken a loan from an SBS-regulated entity. This restriction makes it much more difficult for Financiera Confianza to assess the credit worthiness of individuals who are outside of the formal financial system, thus limiting its universe of potential clients. It is difficult to impossible to serve the unbanked segment.

Second, the financial ratios that Financiera Confianza considers are often difficult to verify and can be either incorrect or nonexistent for small or informal enterprises. Moreover, they reflect past financial performance rather than predict future performance, the key driver of risk for loan repayment. This fact is particularly important in countries with strong economic cycles, as is true of Peru, because two firms may perform well in good times, but it may not be clear which of the two will perform well when economic conditions change. This question is difficult to answer solely looking at historical financial information.

Third, Financiera Confianza's approach requires significant investment in human capital, because of the process by which loan officers establish expertise in assessing an applicant's qualitative risk factors. It is true that such subjective assessments can incorporate a wider set of information for risk assessment and—unlike financial ratios and SBS data analyses—these can be performed for any business (to the extent that the business can provide information requested by an officer). However, these types of assessments are often only as good as the officers performing them. Experienced officers may be finely tuned credit risk analysts who, after years of experience, can incorporate many subtle signals of borrower quality. But, it is expensive for analysts to accumulate this experience. Also, when an individual officer leaves the bank, their experience is lost. A lender's need for personnel training and experience, along with losses from turnover, create limits on how quickly MSME lending can grow. It takes a long time and a large investment to build and retain the required army of experts. Moreover, the high costs per evaluation increase the profitability of large loans relative to small loans, limiting the extent to which banks can serve smaller businesses. While a bank can rely on less experienced loan officers to lower costs, this choice increases the risk of the portfolio.

These three challenges—(i) dependency on borrowing history not available for the unbanked, (ii) uncertainty regarding the ability of screening tools to continue to function in changing economic circumstances, and most importantly, (iii) high operational costs and limits to scale from human-intensive manual processes—are not exclusive to Financiera Confianza or to Peru, but are found across many microfinance institutions globally. Financiera Confianza's approach, while clearly effective in minimizing nonperforming loans, places strong limits on the institution's ability to grow, and the overall financial system's ability to reach informal unbanked enterprise owners. How could financial institutions like Financiera Confianza further evaluate risk for information-scarce MSMEs, particularly informal unbanked firms, and do so while keeping transaction costs low?

Potential Contribution of Psychometrics

Industrial and organizational psychology has been working for decades on a problem with similar characteristics: how to screen people applying for jobs. Firms must decide which individuals to hire, often based on little available information. Moreover, particularly for entry-level positions, firms must evaluate a large number of applicants in a low-cost way. The traditional approach of human-based interviews is costly, limited in scale, and can be biased and

restrict access for certain groups. In response to this challenge, industrial psychology has developed a series of assessments of individuals that are predictive of a person's future success in a job in an objective, quantitative, scalable way. This field, which assesses skills, abilities, personalities, and intelligence, is known as psychometrics.

There is a large established industry using psychometric testing for pre-employment screening. According to a 2001 survey by the American Management Association, 29 percent of employers use psychological assessments of employees for selection and development. Psychometric testing of job applicants is expanding quickly because of its effectiveness. A meta-analysis published by the APA found that tests of general cognitive ability have the highest validity of all selection methods. These tests outperform screening based on personal interviews, biographical data, and past education and experience. The combination of a cognitive ability test and an integrity/honesty test has one of the highest composite validities of any combination of techniques (Schmidt and Hunter, 1998).

To understand the applicability of these tools to financial services, it is relevant to note that job selection is a high-stakes environment where some applicants have very powerful incentives to attempt to "game" the psychometric instruments so as to appear highly qualified for a certain job. Because of these incentives, the assessments used for pre-employment screening have been designed to prevent cheating and have proven highly effective despite an applicant's less honorable attempts (Hogan, Barrett, and Hogan, 2007).

In addition to their practical use in pre-employment screening, psychometric instruments have been extensively used to study the characteristics of good entrepreneurs. These studies typically seek to identify differences between people who become entrepreneurs as opposed to managers or other salaried employees. The literature dates back to McClelland's (1961) seminal work and has considered a wide variety of psychometric tools over the years. Baum, Frese, and Baron (2006) and Chell (2008) provide both systematic literature reviews and meta-analyses.

Given that psychometric tools have proven to be effective screening tools in the human resources field despite the high stakes and incentives to cheat the system, and given that these tools have also been used to distinguish entrepreneurial characteristics among individuals, a few questions arise. Could psychometric instruments help resolve the shortcomings of the traditional approach to credit risk analysis for MSME borrowers? And, could they increase the currently

restricted access to finance among Peru's MSMEs? To answer these questions, we performed a pilot trial of a collection of these tools with Financiera Confianza.

Pilot Trial Implementation

Between May and December 2012, we conducted a pilot study administering a psychometrically enhanced loan application to a representative sample of Financiera Confianza's loan applicants. This section describes the client selection strategy, pilot implementation strategies, and methodologies for gathering this data.

Sample

Literature reviews and new investigations by the Entrepreneurial Finance Lab (EFL) helped to determine the psychometric assessments that were implemented. The EFL tool was administered to 281 applicants randomly selected from a sample of new and existing clients soliciting loans from three Financiera Confianza offices. The bank provided preapproval to the individuals using their screening and scoring system. The amounts disbursed ranged between 500 and 78,000 Soles, with an average loan size of 6,337 Soles and a median loan size of 2,000 Soles. The loans were distributed among four product categories: business products, agricultural products, inclusion products, and personal products. The products primarily differed in their loan sizes and interest rates. Inclusion products are targeted toward the smallest and least formal businesses, offering the smallest loans and highest interest rates. Business and agricultural products are offered to the most formal applicants seeking the largest loans, and they offer the most favorable interest rates. Clients seeking nonbusiness loans and, in some cases, clients whose businesses could not meet the documentation requirements to qualify for business, agricultural, or inclusion loan products, took personal products, which feature less favorable interest rates and loan terms.

Test Implementation

Three strategies were implemented to administer the psychometric assessments to applicants in the pilot study. Each strategy aimed to recruit and contact MSME and individual applicants and administer the EFL tool. In all cases, the process of monitoring adequate data collection was the same.

Recruitment

Strategy 1: Initially, Financiera Confianza loan officers administered the psychometric assessments to preapproved MSME and individual clients before the clients received their loan disbursements.

Recruitment 1: MSME loan officers in three offices (Villa Maria, Villa el Salvador, and Manchay) were trained to administer the psychometric assessments. Each loan officer was then assigned to administer the assessments to a representative sample of clients soliciting the products that formed part of the study.

Strategy 2: Because of the loan officers' struggles to administer the assessments while meeting their other professional responsibilities, the initial volume of assessments was low. Therefore, after 77 days, the recruiting and administration strategy was changed. One specially appointed employee was assigned to coordinate with loan officers to administer the psychometric assessments to pre-approved clients soliciting the products that formed part of the pilot study before the clients received their loan disbursements. The test administrators were then compensated based on the number of assessments they administered to individuals who received business, agricultural, inclusion, or personal loans.

Recruitment 2: Three specially appointed employees, one from each of the offices participating in the study, were trained to administer the psychometric assessments. Each employee then coordinated with loan officers to administer the assessments to clients soliciting the products that formed part of the study.

Strategy 3: Because of high turnover of the specially appointed employees who administered the psychometric assessments, the volume was not sufficient to meet the project's goals. Therefore, after another 76 days, the recruiting and administration strategy was changed again. EFL contracted one employee in each branch who was assigned to coordinate with loan officers to administer the psychometric assessments to preapproved clients before they received loan disbursements. The test administrators were then compensated based on the number of assessments they administered to individuals who received loans.

Recruitment 3: Three individuals, one from each of the offices participating in the study, were hired by EFL and trained to administer the psychometric assessments. Each was then

assigned to coordinate with loan officers in his or her respective branch to administer the assessments to clients soliciting the products that formed part of the study.

Test Administration

EFL loaned each of the offices participating in the project one portable computer, which featured a touch screen. The parties in charge of administering the assessments in each stage of the project were assigned a recruitment target based on the sampling strategy. Thus, officials from branches with the largest portfolios were assigned the largest recruiting targets.

The recruitment procedure began with training loan officers to properly administer the assessment to a random selection of clients who solicited and were approved to receive one of the loan products that formed part of the study. At all times the clients could decide whether or not they wanted to take the assessment. If the client decided to participate, the loan officer administered the test in the office before the client received his or her loan disbursement. Testing typically took place in the waiting area after the loan officer briefly introduced the project and offered basic instructions that explained the testing rules. Administration took place inside the Financiera Confianza office waiting area to cut down on the number of potential distractions the client would face while completing the assessment in their own place of business.

Each participant entered his or her national identification number (DNI) into the assessment at the commencement of the test. Once the test was completed, the results, DNI, and data were stored in the laptop. During test administration, the following adverse situations could take place: (i) the client could struggle to use the computer, forcing the loan officer to enter some or all of the client's answers, thus increasing the potential for answer tampering or (ii) the client could complete the assessment without effort, resulting in haphazard and inconsistent answer selections.

While administering the test, Financiera Confianza monitored the subject to prevent cheating or third party influence on the subject's responses while remaining available to assist with technical issues. During the project, repayment data from 50 applicants was collected using strategy 1, repayment data from 77 applicants was collected using strategy 2, and repayment data from 107 applicants was collected using strategy 3. Loans were ultimately disbursed to 275 applicants that could be matched up to assessment responses in our dataset. SBS data was also available for 254 of those 275 clients.

Throughout the project, EFL and Financiera Confianza also refined and adjusted the implementation strategy based on loan officer and client feedback. For instance, adaptation of the vocabulary used in some of the test questions was required to better fit the targeted demographic groups.

Data Monitoring

Every month the laptops were synchronized in order to send the assessment responses to EFL. Subsequently, EFL sent Financiera Confianza a report with the number of tests completed along with the corresponding DNIs of the individuals that completed each test in order to verify that EFL had received the proper data.

Once the testing phase of the project was completed, EFL prepared and sent to Financiera Confianza a list of subjects that completed the assessment. On June 15, 2013, EFL received the loan performance data from Financiera Confianza, consisting of repayment data from the credit's approval to the date of dataset construction. Because the DNI of the individual who completed the assessment was collected during the assessment's administration, an appropriate matching between the client's repayment data and the test results protected the client's identity. Additionally, EFL obtained the SBS data through May 2013 for all clients who completed the assessment.

Training Sessions

A first meeting to train loan officers from the three participating Financiera Confianza branches took place on May 16, 2012. Subsequent training sessions were conducted on August 1, 2012, for employees administering the assessments between August 1 and October 15, 2012, and on October 23, 2012, for the EFL employee assigned to each of the three offices that administered the assessments between October 24 and December 15, 2012. In each case the training sessions covered three topics: the project's motivation and objective, EFL's psychometric assessment, and the implementation strategy. Given that some of the test questions addressed sensitive personal issues, the training sessions especially highlighted that test results (i) were completely confidential and (ii) would not influence the client's standing with the bank.

It is noteworthy that the person, whether from EFL or Financiera Confianza, administering the assessments played many crucial roles during pilot implementation given his or her proximity to and personal knowledge of the demographic targeted in the pilot implementation. Many adjustments were made to both the implementation strategy and

assessment structure to better tailor it to the target demographic of the pilot study in response to direct feedback from the individuals administering the assessments. Overall, fluent communication and good coordination between EFL, the individuals administering the assessments, and Financiera Confianza's project administrators was critical to ensure a successful pilot implementation.

Results

Cross-Country Comparison

Two groups of countries are available for comparison purposes from other research done by the authors. First, data is available from four major African countries: Kenya, South Africa, Ghana, and Nigeria. This data has the benefit of a large number of tests that are matched to real loan repayment data but has the disadvantage of being culturally dissimilar to the Financiera Confianza data from Peru. The second group of countries is from Latin America: Argentina, Mexico, Peru (from another financial institution in the country), and Costa Rica. While more geographically close and culturally similar, these countries only have test data and are not as yet matched to any loan repayment data.

In addition, both sets of countries featured a different collection of psychometric assessments compared to those implemented for Financiera Confianza. Because of time restrictions during administration of the assessments, the authors had to use an abridged form of the psychometric assessment, thereby excluding a large number of items necessary to calculate some personality dimensions discussed elsewhere (e.g., Klinger et al., 2013). Table 1 summarizes some statistics related to the demographic characteristics of the individuals who completed the psychometric loan application, along with their businesses.

Table 1. Cross-Country Summary Statistics for Comparator Countries

	Gender (% female)	Age (years)	Median business revenues (\$/month)	Business age (years)	Loan size (USD)
Peru (Financiera Confianza)	53%	36	\$1,071	5.9	\$2,300
<i>African Countries with Repayment Data</i>					
Kenya	40%	37	\$6,215	7.5	\$6,750
South Africa	45%	37	\$2,945	3.8	\$6,227
Ghana	34%	40	\$13,206	9.7	\$10,745
Nigeria	23%	40	\$13,102	9.1	\$31,114
<i>Latin American Countries without Repayment Data</i>					
Argentina	16%	49	\$138,392	23	\$50,000
Mexico	31%	46	\$31,142	5.9	\$40,214
Peru (other)	50%	39	\$7,407	8.3	\$9,383
Costa Rica	42%	39	\$2,000	7.2	\$6,470

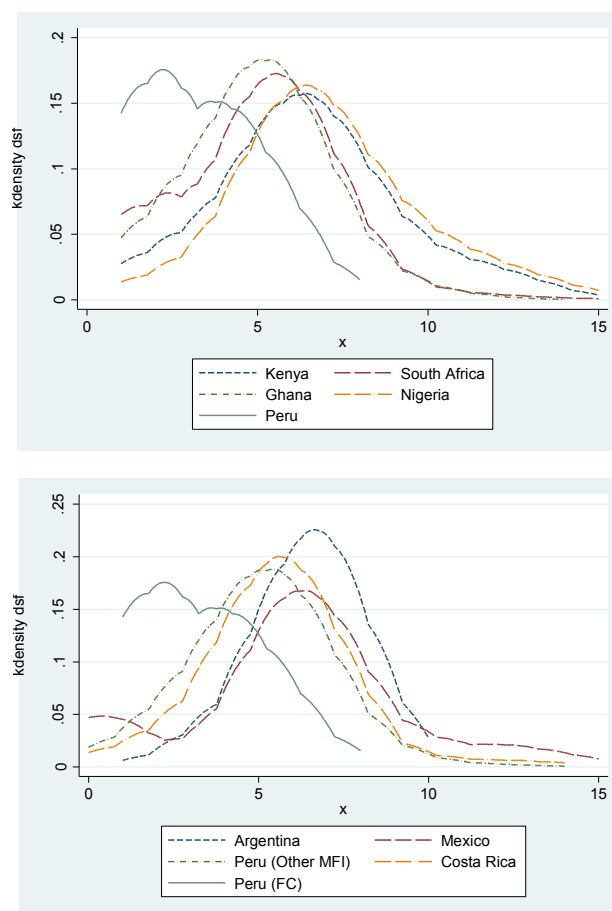
Source: Authors' calculations.

The data from Financiera Confianza is more similar to a microfinance setting than that from the other financial institutions shown in Table 1. Average loan size is one-third the size of those in Kenya, South Africa, Costa Rica, and Peru, which are the closest comparators. Differences in business revenues are even larger. Clients are slightly younger and a larger percentage are female, again consistent with the traditional differences between microfinance and more traditional small business lending by banks. These results suggest that data from other countries will offer an interesting comparison but should be considered in light of the large differences in clients and their businesses, along with the characteristics of the loans they obtained.

In considering comparisons along psychometric questions and content, there is a debate in the literature about the consistency of psychometric instruments across countries and cultures. Strong cultural differences between countries are often reflected in different average answers and scores on a wide variety of psychometric tools (e.g., the World Values Survey at www.worldvaluessurvey.org). Yet evidence exists that across countries as diverse as Canada, China, Israel, Japan, the Philippines, Poland, and South Africa, major psychometric constructs, such as the “Big Five” personality structure (the predominant personality model in psychology), remain a consistent construct (McCrae et al., 1998).

Unfortunately, the “Big Five” were not measured in the Financiera Confianza study because of time limitations during the loan application period. But one dimension that was included in the Financiera Confianza assessment as well as the assessments used by the institutions listed above is fluid intelligence. Intelligence is often thought of as a person’s general ability to think and act rationally. The APA formed a task force on intelligence research and defined it as the “ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, and to overcome obstacles by taking thought” (Neisser, Boodoo, and Bouchard, 1996). Fluid intelligence describes the ability, regardless of prior skills or experience, to grasp relationships between abstract concepts. Fluid intelligence is usually measured by tests of problem-solving ability, pattern recognition, and working memory (Snow and Yalow, 1982). In this case, it was measured using a digit span recall test, which is a traditional subtest in the Wechsler Adult Intelligence Scale.

Figure 7. Digit Span Recall Results by Country



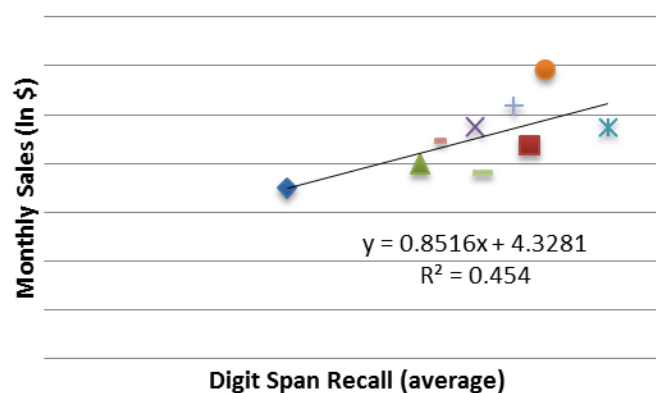
Source: Authors’ calculations.

Compared to the African countries, the Financiera Confianza loan applicants scored significantly lower on the digit span recall test. They correctly recalled an average of only 3 digits, compared to 5 in Ghana and South Africa, 6 in Kenya, and 7 in Nigeria. These are not just cultural or regional differences, as the average of 3 is also lower than Mexico, Argentina, and Costa Rica (6), and even lower than the other financial institutions' applicants from Peru (5).

Intelligence has been repeatedly shown to be the best single predictor of job performance in both entry-level and advanced vocations (Hunter, 1986; Ree, Earles, and Teachout, 1994). Because pattern recognition plays a strong role in profitable entrepreneurship (Baron, 2006) and because successful entrepreneurs must be a “jack of all trades” rather than a specialist (Lazear, 2005), fluid intelligence could play an even stronger role than it plays in general vocational performance.

There is a clear pattern across countries between intelligence as measured in a digit span recall exercise and sales levels. This can be most easily observed in a plot of average digit span recall and the natural log of median monthly sales for each country (Figure 8). This pattern is maintained looking within countries as well as across them. Performing a regression of revenues on digit span with bank fixed effects (which isolates the relationship within countries rather than across them) continues to result in a positive and highly statistically significant relationship, though economically only of marginal significance (the difference between 3 and 5 additional digits recalled correctly was associated with an additional 10 percent per month in sales for Financiera Confianza clients).

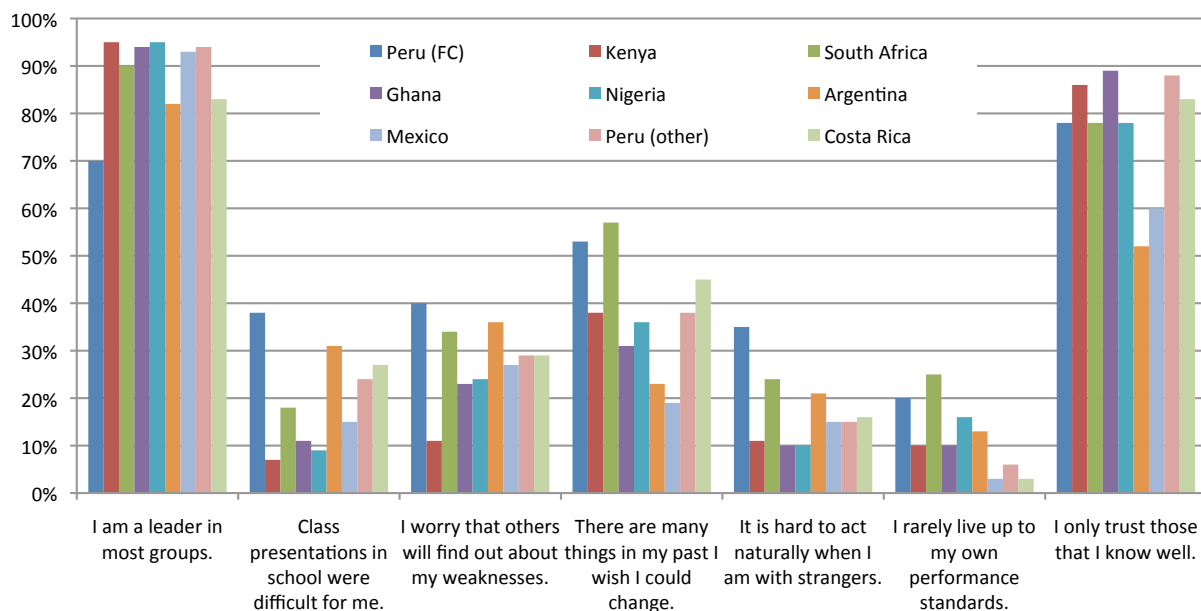
Figure 8. Digit Span Recall Results and Monthly Sales by Country



Source: Authors' calculations.

In addition to a test of fluid intelligence, the test administrators posed various psychometric questions to applicants. Because of time constraints it was not possible to implement a full Big Five personality model during the application process (Barrick and Mount, 1991); however, the administrators asked a subset of questions from a leading personality assessment. For confidentiality reasons, these items cannot all be shown here; but for illustrative purposes, Figure 9 presents seven selected items that address a range of personality facets and subfacets.

Figure 9. Percentage of the Sample that Agrees with Each Statement by Country



Source: Authors' calculations.

As with the digit span recall results, Figure 9 shows that there are significant differences between countries in the sample. There could be many causes for these differences. One is the cultural difference between countries. Responses to personality items can vary significantly across countries, as words like “act naturally” can mean very different things depending on an individual’s place of origin. If we were to assume that LAC countries are more similar to one another than to African countries, and similarly for African countries as compared to LAC countries, then we would expect to see some regional trends in the response differences above.

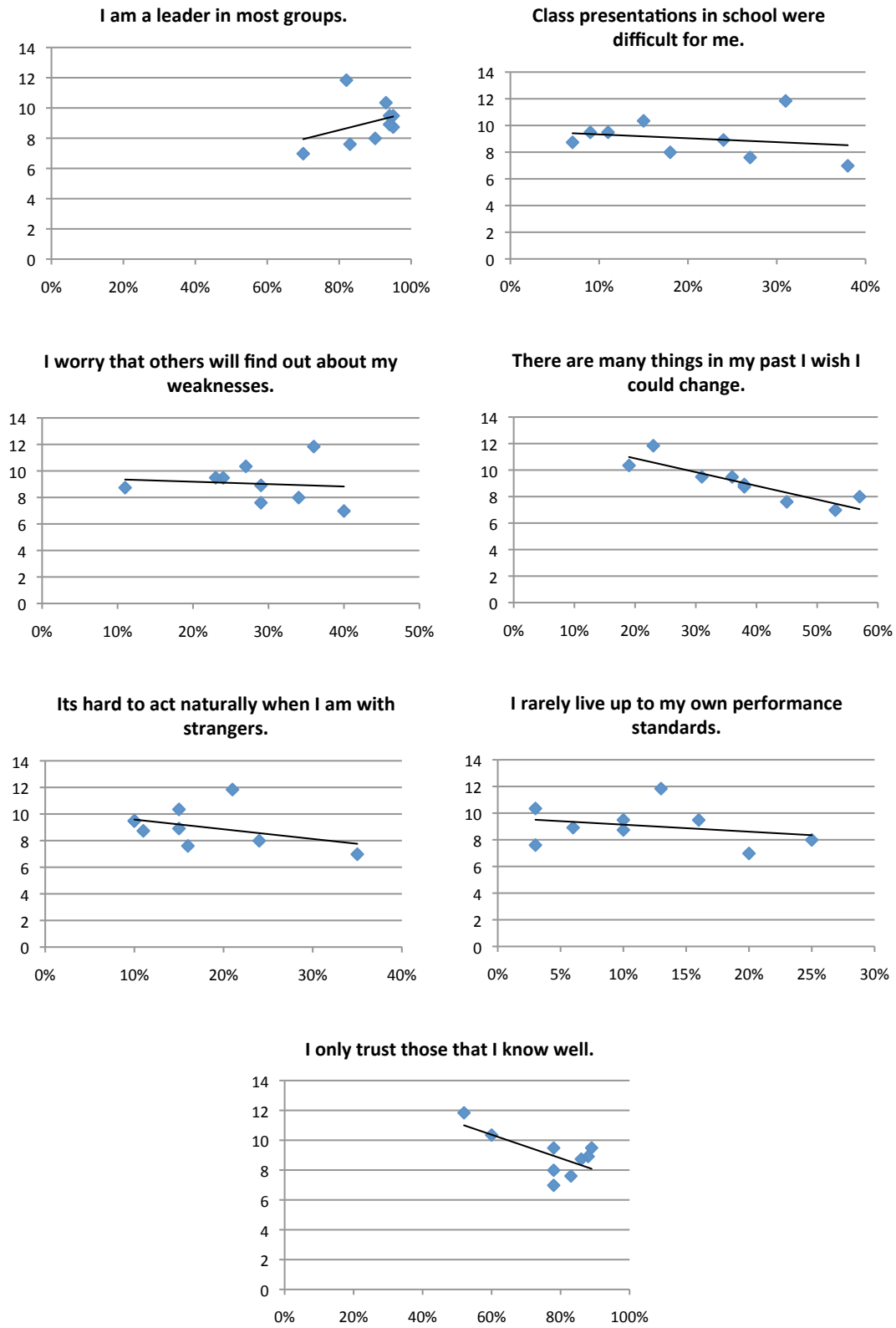
While there are no dominant regional patterns, there is a notable tendency to agree less with the statement “I am a leader in most groups” and a tendency to agree more with the statements “Class presentations in school were difficult for me,” “I worry that others will find out my weaknesses,” and “it’s hard to act naturally when I am with strangers” for LAC countries compared to African countries. Interestingly, in almost all of these cases, South Africa is more similar to the LAC countries than the other African countries. This could potentially relate to the fact that income per capita in South Africa is also more similar to the LAC sample than the other African countries in the sample, though such a relationship is only suggestive and cannot be established with the above data.

Another source of differences between countries is the fact that the financial institutions within each country making up this sample are dealing with different business segments, some of which are microenterprises and others small businesses. As with intelligence, there are some cross-country patterns that seem to relate to business size. Figure 10 shows the equivalent plots of the percentage of entrepreneurs who agree with the statement against business sales.

The patterns are not as strong as with the digit span recall test. Some are stronger, such as “I only trust those I know well” and “There are many things I wish I could change about my past”, while others are very weak, such as “I rarely live up to my own performance standards” and “Class presentations in school were difficult for me.”

As with digit span, the cross-country patterns are interesting and could partially be related to the relationship with real business outcomes of importance, such as sales or default risk. But for psychometrics to be useful for entrepreneurial screening, it is not even necessary that the measured traits have equal meanings and distributions across countries. It is more important that the same dimensions predict risk within each context. For example, in determining whether height is a good predictor of success at basketball, it does not matter that average height is lower in one country and higher in another. What matters is that height is correlated with improved performance at basketball so that basketball abilities can be predicted by comparing players’ heights within each country.

Figure 10. Percentage of the Sample that Agrees with Each Statement by Country vs. Business Sales



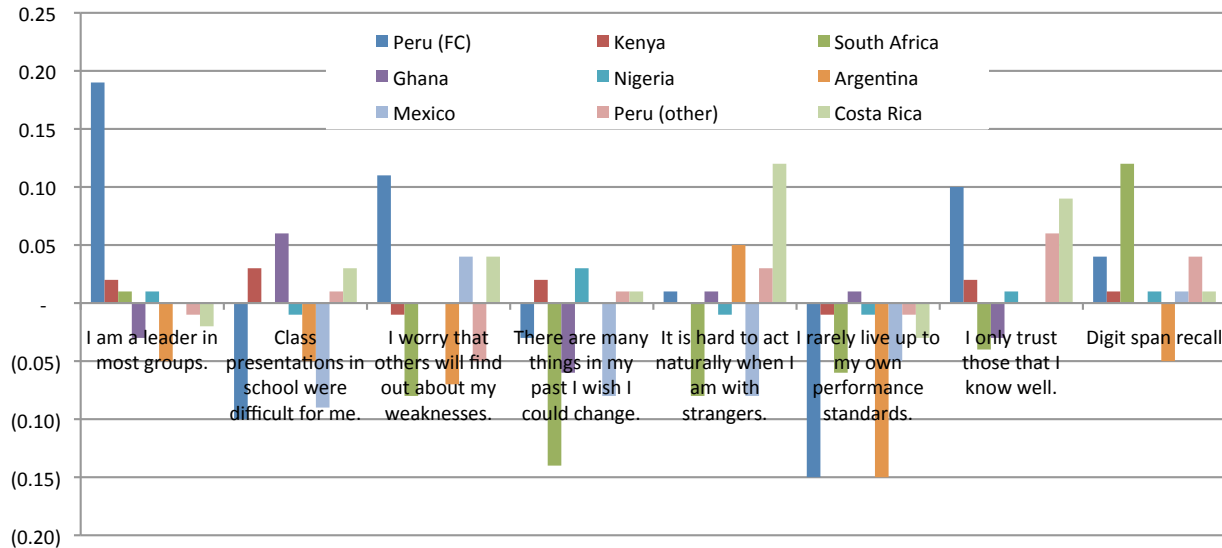
Source: Authors' calculations.

An emerging study by Patterson et al. (2012) examines an issue that is similar but related more to managerial effectiveness than to entrepreneurial success. The authors studied whether the importance of key skills and attributes such as taking action, making decisions, following through, developing relationships, and having drive and ambition vary by national culture (divergence) or are universally similar (convergence). Despite some research showing cultural differences in management styles and behaviors, proponents of the convergence hypothesis suggest that globalization and the prevalence of multinational corporations can cause cultural norms and values to converge. The authors tested managers across 30 countries and found support for the convergence hypothesis. Indeed, the importance of the aforementioned skills and attributes for successful management are not shaped by culture but are similar across all 30 countries.

To the extent that managerial and entrepreneurial tasks are similar, these findings support the hypothesis that the particular traits affecting entrepreneurial performance (and thereby, credit risk) could also converge across cultures, allowing similar instruments to be used in multiple countries. Just as companies are increasingly working across borders, small businesses and the firms that train and invest in them are also increasingly transnational. Non-governmental organizations (NGOs) such as TechnoServe and Endeavor identify, select, and train entrepreneurs across nearly all regions of the world. The emerging industry of “impact investing,” which focuses on bottom-of-pyramid businesses in developing countries and their environmental and social impacts, is also applying selection methodologies and sending investment officers across a wide variety of countries (e.g., members of the Aspen Institute’s Aspen Network of Development Entrepreneurs). The ability of individuals to source deals and select entrepreneurs across multiple countries and cultures suggests that at least some of the drivers of entrepreneurial ability are similar.

Scoring models are also used to order individual applicants within a particular country or for a particular bank, market segment, or product. Therefore, whether the findings of a psychometric-based application can be generalized across countries and cultures merely requires that the same questions be able to distinguish entrepreneurial success or default risk *within* different groups. Figure 11 shows the correlation between each of the example personality items discussed above, along with digit span recall and business revenues.

Figure 11. Correlation between Specific Personality Items, Intelligence, and Business Revenues by Country

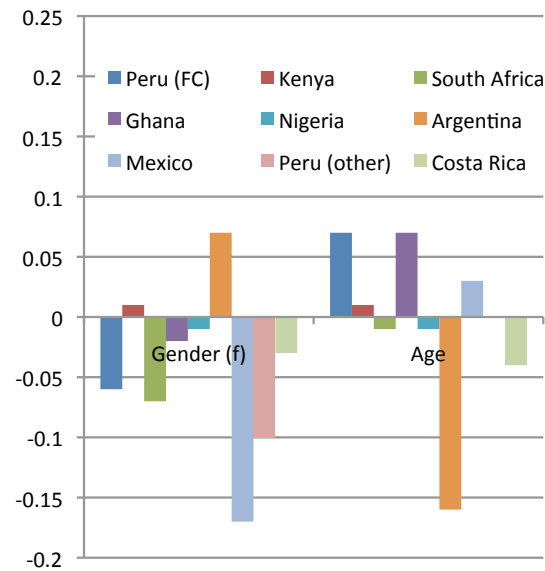


Source: Authors' calculations.

Looking within country, the picture is far less clear. Among these seven personality items, as well as digit span, there are very few consistent patterns. Digit span recall is related to higher sales in almost all countries other than Argentina, and agreeing with the statement “I rarely live up to my own performance standards” is negatively related to business sales in all countries other than Ghana. The other items do not have consistent relationships with business revenues across all countries, and the differences do not relate in any obvious way to regional differences.

Does this imply that psychometric questions could not be a useful screening tool for identifying high-revenue entrepreneurs? Figure 12 shows that demographic characteristics, which are also often used in entrepreneurial selection, such as for credit scores of unbanked entrepreneurs, are similarly inconsistent across countries when considering business revenues.

Figure 12. Correlation of Demographics with Business Revenues



Source: Authors' calculations.

More generally, it is important to note that the psychometric items above are only a subset of the entire set of questions asked on the Financiera Confianza psychometrically enhanced credit application. Thus, there could be a larger collection of items that, similar to digit span and “I rarely live up to my own performance standards,” have consistent relationships across countries and could form a cross-country generic screening tool. But the results above suggest that there are significant differences between countries as well. Either this cross-country variance is simply statistical noise (which can easily be checked with out-of-sample testing by country), or the stronger relationships between psychometric (and demographic) items are particular to certain countries and market segments, and must be customized to achieve their true predictive power.

Predicting Credit Risk

The results above relate some sections of the psychometric credit application to business revenues. Predicting revenues could be of interest to organizations seeking to identify high achieving entrepreneurs for promotion, assistance, or direct investment by venture capitalists, angel investors, and others. But when considering access to external funding, angel and venture capital investment is extremely uncommon. Even in the United States, for example, venture capital firms finance less than 0.03 percent of all new businesses founded in the country each

year (Shane, 2010). It is loans from commercial banks that are the leading source of external debt finance for new businesses in the United States (Shane, 2010) and, given that emerging markets have nowhere near the level of venture capital and angel investor activity that is found in the United States, commercial bank finance for entrepreneurship in those emerging markets is likely even more important.

Therefore, as a screening tool to promote more lending by commercial banks and microfinance institutions, prediction of default rather than revenues matters more. This could be somewhat related to predicting sales, as business owners need to generate sufficient cash in their businesses to be able to repay loans. However, loan repayment depends both on ability and willingness to repay and, even when speaking only of ability to repay, lenders do not prefer entrepreneurs who have higher, but potentially more volatile, business performance. With their upside capped at the interest rate, lenders only care about an applicant's ability to consistently earn enough to make loan payments.

For all 275 Financiera Confianza clients in the pilot study, we obtained loan repayment data from two sources. First we looked at loan repayment performance during the first six months of their loan from Financiera Confianza. Of those 275 clients, 35 (12 percent) went into 30 days of arrears or more, meaning they missed one payment and passed the due date of the next payment, meaning they entered the status of two payments owed to the bank. Of those, nine (3.0 percent) went 60 days in arrears and seven (2.6 percent) went into 90 days or more of arrears.

In addition to loan repayment performance from Financiera Confianza, we also obtained credit usage reports on 254 of the 275 clients from the SBS. This data allowed us to track the applicants' use of loans across the entire Peruvian financial system, including payments to other credit providers such as utilities. Among those 254 individuals, 101 (40 percent) entered into at least 30 days arrears to a credit provider, 42 (17 percent) of greater than 60 days, and 29 (11 percent) of greater than 90 days.

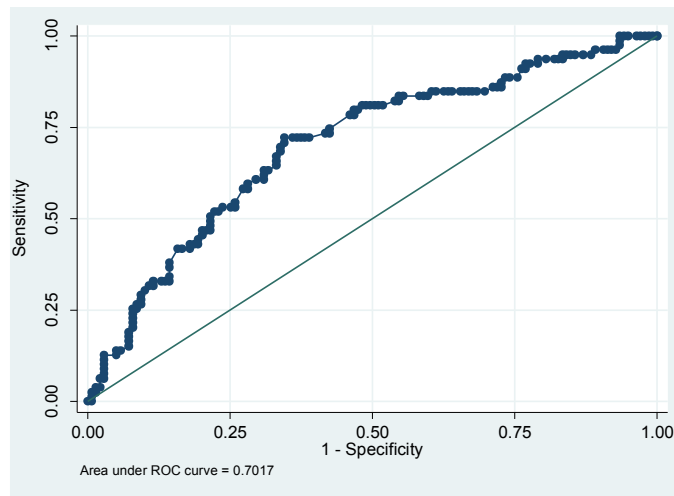
This gives us a wide variety of loan repayment data to consider. The goal of any credit-screening tool is to identify bad applicants ("bads") that the lender does not want as clients. Arrears of 30 days or more has the benefit of providing us a larger number of bads clients for statistical testing of our credit scores, but this is a weak demonstration of bad behavior, as a large percentage of clients could miss one payment due to carelessness, travel, or temporary cash

flow mismanagement, yet recover and fully repay their loan. In most banks, more than half of clients who enter into 30 days of arrears recover and repay fully and are granted subsequent loans. So at 30 days there are more bads but this is a weaker signal of being bad. Increasing to 60 or 90 days becomes a stricter definition, where a much higher percentage of borrowers continue to default, but in our sample is a smaller number of bads and therefore could be less meaningful statistically. Similarly, the system-wide arrears status from the SBS provides a larger number of bads but may also be a weaker definition of bad. Indeed, the spouse of one of the authors of this paper would count as a bad applicant because of a disputed US\$5 service fee on a cancelled credit card, though she has never missed a loan repayment. We therefore show model performance metrics for all combinations of sources and levels of arrears so that these trade-offs can be considered.

For it to be effective in practice, the optimal responses and weightings of a credit application scorecard must be confidential, so we will not show the relationships between any individual items or questions and default, only the overall score's predictive power. The model's predictive power is measured by the area under the receiver operating characteristic curve (area under the curve, or AUC), a common metric in credit scoring. The other common metric in the industry is the Gini coefficient, which is simply two times AUC minus one. For more details on how to calculate AUCs and Gini coefficients, see the appendix.

First, we examine the Financiera Confianza data directly. The sample size is very small, so we have to consider the broadest definition of bad (30 days or more of arrears across all of the financial system). We conducted a forward stepwise logistic regression with a 10 percent cut-off in the stepwise procedure, considering all 66 features. The resulting model has an AUC of 0.7 (Gini of 40 percent) and is shown in Figure 13.

Figure 13. Model Built on Financiera Confianza Data

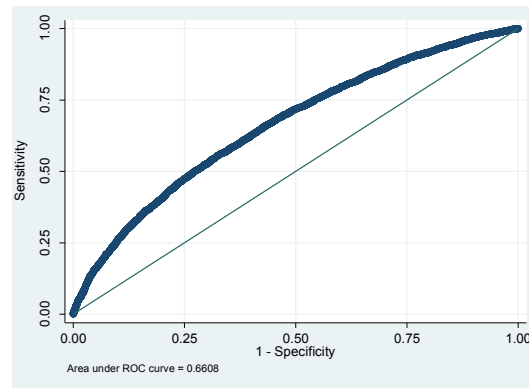


Source: Authors' calculations.

Though this model has a very high AUC, it is surely over-fit to the data. Indeed, with so many potential covariates that could enter the stepwise regression, data consisting of pure noise could generate this result looking in-sample. It is best practice to always look only at a randomly selected hold-out sample to really evaluate the predictive power of a credit-scoring model once implemented. But the Financiera Confianza sample is too small to do so.

Therefore, we pursued a much more difficult test of a psychometrically enhanced credit application: building the model on other countries and applying it to Financiera Confianza. We built a model on all of those items included in the Financiera Confianza application that were also included in the applications from those countries shown above that have loan performance data (Kenya, South Africa, Ghana, and Nigeria), performing the equivalent stepwise procedure. A total of 20 features entered into this model, of which two were demographic and the remaining 18 psychometric. The cross-country results in the previous section indicate significant differences in applicant population and cultures between these African countries and Financiera Confianza, making this an extremely high bar to pass. We built a model on African small business data and applied it to Peruvian microenterprises without using any Peruvian data for model construction. If even moderate predictive power remains, this is a powerful signal. Within the African data, this model achieves an AUC of 0.66, which is a Gini coefficient of 36 percent (Figure 14).

Figure 14. ROC Curve - African Model Applied to African Data



Source: Authors' calculations.

Table 2 shows the AUC and Gini values for all three potential bad levels for Financiera Confianza loan status only and default status across the entire financial system.

Table 2. AUC and Gini of African Model Applied to Financiera Confianza

	30 days or more	60 days or more	90 days or more
Financiera Confianza loan history	AUC: 0.56 Gini: 12%	AUC: 0.58 Gini: 16%	AUC: 0.53 Gini: 6%
SBS loan history	AUC: 0.58 Gini: 16%	AUC: 0.59 Gini: 18%	AUC: 0.60 Gini: 20%

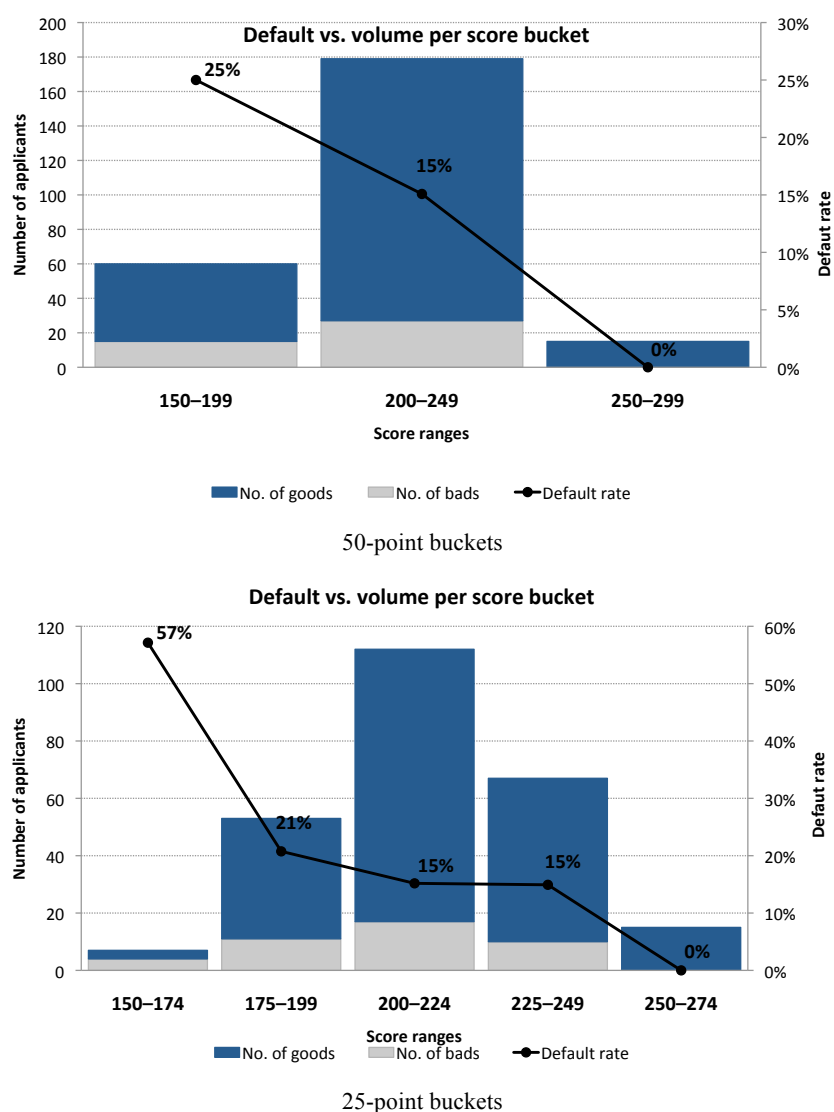
Source: Authors' calculations.

Since Gini and AUC levels are sensitive to the population default rate and their value also depends on the costs of lending to a bad versus rejecting a good borrower, there are no simple benchmarks on what is a good Gini. It also significantly depends on what kind of data is available. Behavioral scorecards (predicting repayment next month using actual loan repayment data of all previous months) have much higher Ginis than application scorecards, particularly for unbanked and new-to-bank applicants (as is the case for Financiera Confianza applicants). But if psychometric content can provide a 20 percent contribution to Gini for difficult to evaluate micro and small business owners, this could dramatically improve access to finance for MSMEs and open large profitable new market segments for banks. In the results above, this level is nearly achieved using a model built on data from another continent, on relatively larger small businesses. True predictive potential of this psychometrically enhanced application scorecard that was customized to Financiera Confianza data is surely smaller than the 40 percent calculated

in-sample, but surely larger than these results that are applied from a set of significantly different African countries. A Gini coefficient of 20 to 40 percent approximates to a probability of default that is three to four times higher for those below any hypothetical accept/reject cut-off selected based on the score to those above that cut-off.

To show this same result another way, Figure 15 charts default rates using a bad60 from the SBS definition of bad against the model built from African data (AUC=0.59). We show numbers of clients and default rate by score range, both in 50-point buckets and in 25-point buckets.

Figure 15. African Model on SBS 60Days+ Arrears for Financiera Confianza

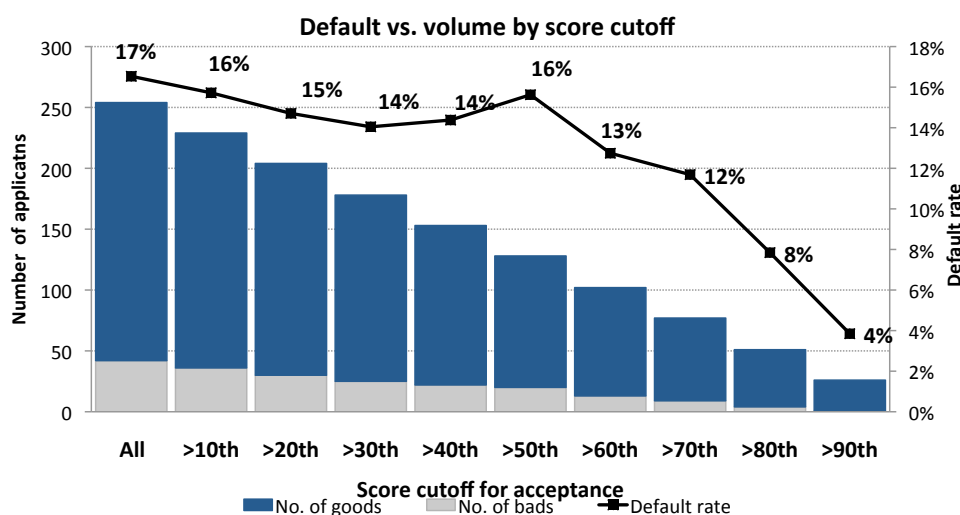


Source: Authors' calculations.

These charts show that the bad rate for those with scores under 200 is 66 percent higher than for those between 200 and 250, while there were no bad cases for applicants scoring above 250 on the psychometric application score (though those are a small number). Showing the same results by a more disaggregated 25-point bucket definition, we see similar results. While there is not as much risk differentiation in the middle of the distribution (200 to 250 points), we see that the model is able to identify high risk and low risk populations around this middle.

This score could be used for other decisions, such as setting interest rates or collateral requirements, or to fast-track low-risk applications and/or flag high-risk applications for review. However, the typical use of these scorecards is to set an accept/reject cut-off. The optimal cut-off depends on financial analysis of the costs of accepting bads and the foregone profit of rejecting goods. But in general terms we can show the resulting bad rate for all accepted clients at ever increasing cut-offs (Figure 16).

Figure 16. African Model on SBS 60Days+ Arrears for Financiera Confianza vs. Volume by Score Cut-off (by decile)



Source: Authors' calculations.

Again, this African model does not monotonically sort risk within the middle-scoring cohorts, as the default rate rises slightly at some points in the middle of the distribution. This is anticipated simply because of the relatively small sample size and number of bads, which is expected to bring greater variability that would be smoothed out with more data. But even more importantly, the model itself would be expected to improve its risk sorting power as it was

customized to Latin America and Peru, incorporating those local and regional differences that were also seen between psychometric content and revenues in the previous section. Nevertheless, even with this crude approximation using only African data, the overall bad rate in this Financiera Confianza sample could be reduced to 15 percent from 17 percent by rejecting the bottom 20 percent of applicants. Those rejects have a 24 percent bad rate, which is unlikely to be profitable depending on the interest rate. Similarly, the bad rate could be further reduced to 14 percent by rejecting the bottom 30 percent of applicants.

It is important to note that these results are only for existing clients who, despite the barriers to financial access mentioned at the outset, were able to receive funding. Reducing risk and operational expenses in serving this segment is a value added of a psychometric application scorecard, but the even larger benefit is the ability to use this tool to access new clients who normally would have been rejected. We have seen that very limited credit depth in Peru is concentrated among smaller businesses, independent workers, and informal businesses, largely because information about their risk is not available. But this psychometric data is available from every potential applicant and could allow institutions such as Financiera Confianza to further penetrate this high-potential underserved segment.

Conclusions and Further Directions

Peru has a unique situation of low financial depth combined significant competition in the financial sector and of high rates of access to finance on favorable terms for more formalized firms combined with expensive or impossible access to finance for smaller, informal unbanked entrepreneurs. This could be related to the limitations of current financial screening tools used by banks and microfinance institutions like Financiera Confianza, namely a dependency on collateral, borrowing history, and subjective manual evaluations by credit officers that are expensive to train and who rapidly change jobs. Enhancing risk analysis using psychometrics is a potential solution to this constraint since it could provide more information for risk analysis that is available from all firms, even informal unbanked enterprises, at a low cost and with high potential for scale. We evaluated such a tool on a sample of Financiera Confianza loan applicants.

The results of this pilot study show that, unsurprisingly, citizens of different countries clearly have personality patterns and psychological characteristics that differ based on divergent

cultures, languages, and values. Yet despite this diversity, dimensions that are related to entrepreneurial performance and credit risk have some commonalities across countries. A psychometrically enhanced scorecard built on Financiera Confianza's data achieves a Gini of 40 percent, while that built on African data achieves a Gini of up to 20 percent. Moreover, because these tests were applied on approved applicants, this discriminating power is over and above traditional screening tools that use some of the subjective assessments and financial ratios described above.

Therefore, psychometric-based scorecards would allow Financiera Confianza to lower its current levels of risk and operational expense among currently served markets segments. But even more exciting, both from a development perspective and from a business opportunity perspective, is the potential to reach MSME owners that are traditionally rejected due to a lack of business and credit history, collateral, and co-signers. Reaching MSMEs is feasible with a psychometric application because the data can be collected from all applicants and has significant potential, particularly because the results shown in this paper are for a model developed with African data, which has not yet been adapted to the local context in any way. Over time, as psychometric content can be fused onto loan applications and the underlying models customized to cultural and market particularities, this type of tool could help resolve the challenge of reaching more informal microenterprises in Peru, which is a priority given the significant amount of competition and financial access enjoyed by more formal small businesses.

References

- Baron, R. 2006. "Opportunity Recognition as Pattern Recognition: How Entrepreneurs 'Connect the Dots' to Identify new Business Opportunities." *Academy of Management Perspectives* 20: 104–19.
- Barrick, M., and Mount, M. 1991. "The Big Five Personality Dimensions and Job Performance: A Meta-Analysis." *Personnel Psychology* 44: 1–26.
- Baum, R., Frese, M., and Baron, R. 2006. *The Psychology of Entrepreneurship*. New York, NY: Psychology Press.
- Chell, E. 2008. *The Entrepreneurial Personality: A Social Construction*. London, UK: Routledge.
- CODEMYPE (Consejo Nacional para el Desarrollo de la Micro y Pequeña Empresa). 2011. "Estadísticas." Lima, Peru: Dirección General de Mype y Cooperativas, Ministeria de la Producción. Available at <http://www.produce.gob.pe/remype/data/mype2010.pdf>.
- EIU (The Economist Intelligence Unit). 2009. "Global Microscope on the Microfinance Business Environment." Commissioned by the IDB and International Finance Corporation. *The Economist*. Available at: http://www1.ifc.org/wps/wcm/connect/98651e004a181499b9e0fddd29332b51/Microfinance_ENG_WEB_Sept%2B25.pdf?MOD=AJPERES.
- Ferraro, C. 2011. "Eliminando barreras: el financiamiento a las Pymes en America Latina." Agencia Española de Cooperación Internacional para el Desarrollo (AECID) for the Programa de Cooperación CEPAL-AECID. Santiago, Chile: United Nations. Available at: http://www.eclac.org/ddpe/publicaciones/xml/0/45220/Financiamiento_a_pymes.pdf.
- Hogan, J., Barrett, P., and Hogan, R. 2007. "Personality Measurement, Faking, and Employment Selection." *Journal of Applied Psychology* 92(5): 1270–85.
- Hunter, J. 1986. "Cognitive Ability, Cognitive Aptitudes, Job Knowledge, and Job Performance." *Journal of Vocational Behavior* 29(3): 340–62.
- INEI (Instituto Nacional de Estadística e Informática). 2011. "Encuesta de micro y pequeña empresa 2010." Lima, Peru: Dirección Nacional de Censos y Encuestas del Instituto Nacional de Estadística e Informática. Available at: <http://www1.inei.gob.pe/BiblioINEIPub/BancoPub/Est/Lib0980/libro.pdf>.
- Klinger, B., Castro, L., Szenkman, P., and Khwaja, A. 2013. "Unlocking SME Finance in Argentina with Psychometrics." IDB Technical Note No. IDB-TN-532. Washington, DC: Inter-American Development Bank.
- Lazear, E. 2005. "Entrepreneurship." *Journal of Labor Economics* 23(4): 649–80.
- McClelland, D. 1961. *The Achieving Society*. Princeton, NJ: Van Nostrand.

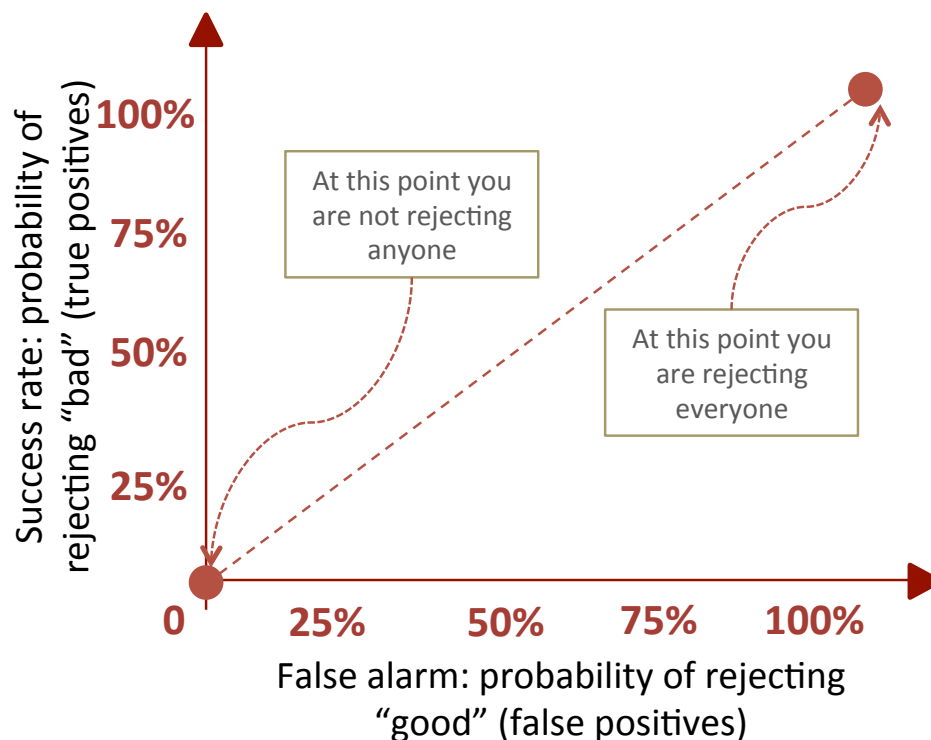
- McCrae, R. R., and Costa Jr., P. T. 1999. "A Five-Factor Theory of Personality." In O. John, R. Robins, and L. Pervin (eds.) *Handbook of Personality: Theory and Research 2*. New York, NY: The Guilford Press.
- Neisser, U., Boodoo, G., and Bouchard, T. 1996. "Intelligence: Knowns and Unknowns." *American Psychologist* 51(2): 77–101.
- Quispe, Z., León, D., and Contreras, A. 2012. "El exitos desarrollo de microfinanzas en el Perú." Lima, Peru: Central Reserve Bank of Peru. Available at: <http://www.bcrp.gob.pe/docs/Publicaciones/Revista-Moneda/moneda-151/moneda-151-03.pdf>.
- Oxford Business Group. 2012. *The Report, Peru 2012*. London, UK: Oxford Business Group.
- Patterson, A., Gentry, W. A., Stawiski, S. A., Gilmore, D. C., and Sparks, T. E. 2012. "A 30-Country Multilevel Analysis of Managerial Skillsets Important for Success." Paper presented at the 27th annual meeting of the Society for Industrial and Organizational Psychology, San Diego, California.
- Ree, M., Earles, J., and Teachout, M. 1994. "Predicting Job Performance: Not Much More than g." *Journal of Applied Psychology* 79(4): 518–24.
- Schmidt, F., and Hunter, J. 1998. "The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of Research Findings." *Psychological Bulletin* 124(2): 262–74.
- Shane, S. A. 2008. *The Illusions of Entrepreneurship: The Costly Myths that Entrepreneurs, Investors, and Policy Makers Live By*. New Haven, CT: Yale University Press.
- Snow, R., and Yalow, E. 1982. *Handbook of Human Intelligence*. London, UK: Cambridge University Press.
- World Bank. 2010a. "World Bank Enterprise Survey, Peru 2010." Washington, DC: The World Bank. Available at <http://www.enterprisesurveys.org/Data/ExploreEconomies/2010/peru>
- . 2010b. "World Bank Enterprise – Informal Survey. Peru – Informal Survey 2010." Washington, DC: The World Bank. Available at <https://www.enterprisesurveys.org/Portal/Login.aspx?ReturnUrl=%2fportal%2felibrary.aspx%3flibid%3d14&libid=14>
- . 2012a. "World Bank Data, Peru 2012." Washington, DC: The World Bank. Available at <http://data.worldbank.org/country/peru>.
- . 2012b. "World Development Indicators Data, 2012." Washington, DC: The World Bank. Available at <http://data.worldbank.org/data-catalog/world-development-indicators/wdi-2012>.

Appendix: Graphing and Interpreting AUC and Gini Coefficient

A common set of metrics is used to assess credit scores. It is important to remember that any credit score does not trigger a decision to accept or reject an applicant; assessment is a continuous relative measure of risk, and lenders can make an accept/reject decision based on any score cut-off point they select. Metrics to assess the predictive power of credit scores therefore evaluate the score's ability to sort applicants by credit risk rather than by imposing any single cut-off. If a score is closely related to default, then participants with a low score should be much more likely to default than those with a high score. Credit scores with little value for directing lending do not separate high-risk applicants from low-risk applicants, and both categories tend to be evenly distributed across the score's spectrum.

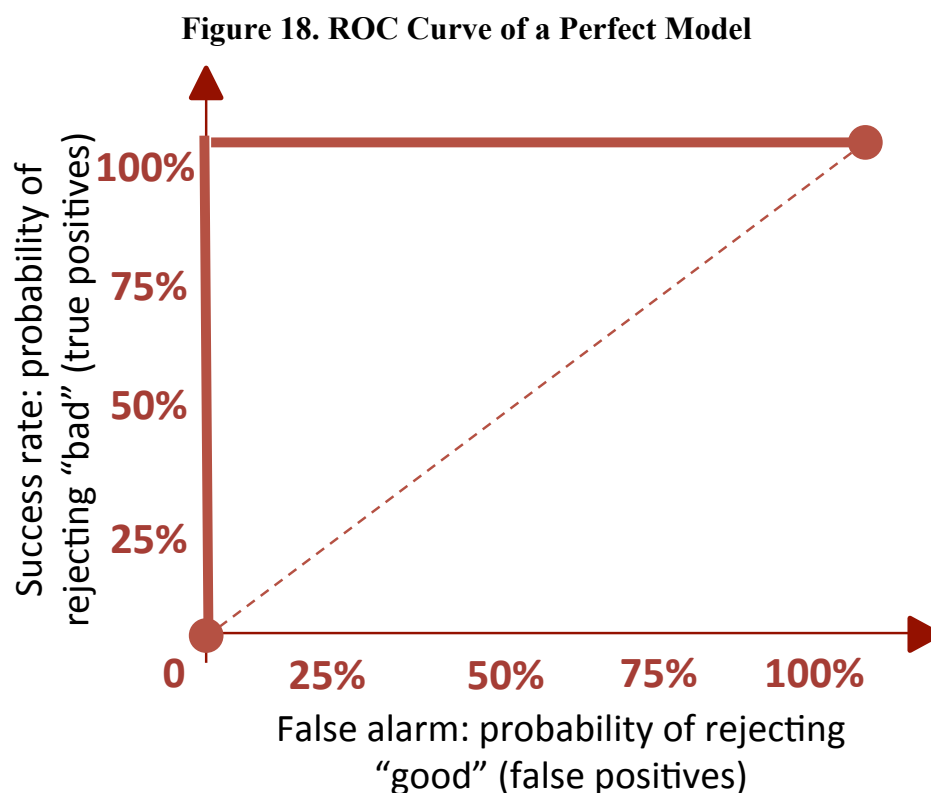
This ability of a model to sort applicants based on level of default risk is typically illustrated by a receiver-operating characteristic curve, or ROC curve. This curve plots on the x-axis the percentages of “goods” (non-defaulters) below any particular score level, while the y-axis shows the percentage of “bads” (defaulters) below that score. Any credit scoring model represents a curve on this graph, with each point on the curve showing the impact of a potential cut-off score.

Figure 17. Building an ROC Curve



Source: Authors' calculations.

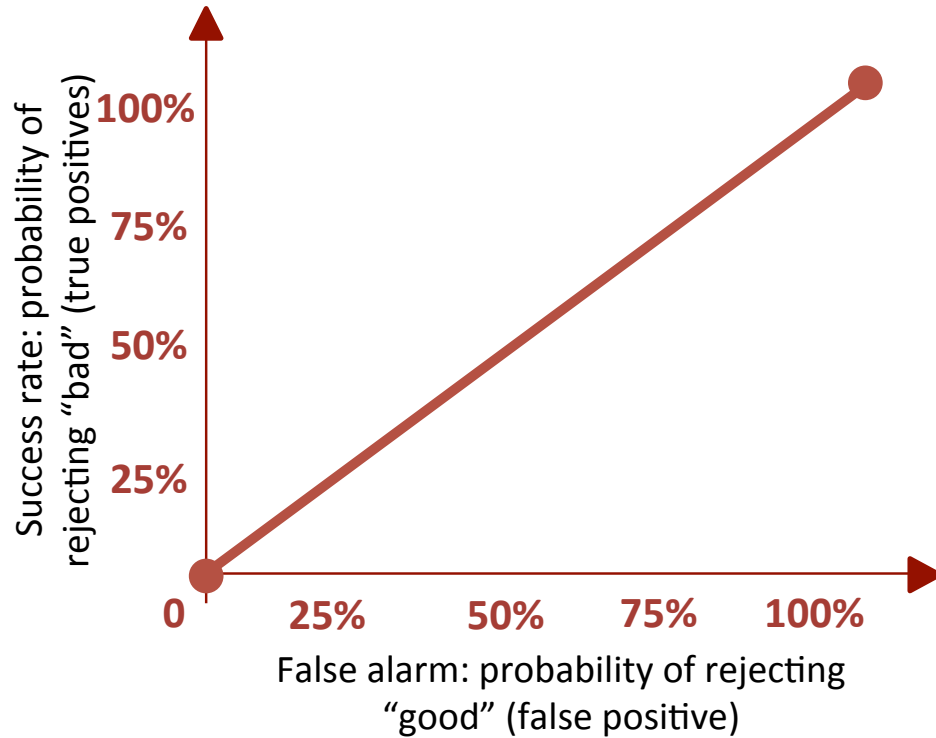
A perfectly predictive credit model would assign the lowest score to all defaulting clients, and therefore, in Figure 18, if you started rejecting applicants with the lowest score you would only reject defaulters, producing a move up the y-axis while the x-axis remains at 0 percent. Only after raising the rejection score cut-off to the point that 100 percent of the bads were rejected (top left corner) would raising the rejection score cut-off start rejecting goods, moving from the top left corner horizontally along the y-axis until it reaches the maximum score so that all applicants would be rejected (top right). So the ROC curve for a perfectly predictive credit score would appear at the top left side of the figure.



Source: Authors' calculations.

Conversely, the ROC curve for a credit score containing no predictive power would fail to distinguish bads from goods at any level but would be like flipping a coin. This means good and bad credit risks would be evenly distributed across all scores. So in Figure 19, beginning from the lowest score and increasing the rejection score cut-off would lead to a rejection of both goods and bads in equal proportion. In other words, the ROC curve would be a straight diagonal line.

Figure 19. ROC Curve of a Random Model



Source: Authors' calculations.

Most scoring models fall between these two extremes. A better credit-scoring model would more resemble the 90-degree angle in Figure 18 than the diagonal line in Figure 19. When the line bows up and to the left, the model gives lower scores to a greater proportion of bads and higher scores to a greater proportion of goods.

In summary, the credit scoring industry typically describes a model's power using the area under the ROC curve (AUROC or often shortened to AUC). The higher an AUROC, the better the model. The perfectly predictive model above (Figure 18) has an AUROC of 1, while the useless model (Figure 19) has an AUROC of 0.5. A common transformation of the AUROC is a Gini coefficient, which is simply $(2 \times \text{AUROC}) - 1$. While comparing AUROC or Gini coefficients across different samples is not recommended because they are sensitive to absolute level of default, these metrics are useful for comparing the performance of different models using the same data.