Economic Impacts of Payment Reporting Participation in Latin America

Michael Turner, Ph. D.
Robin Varghese, Ph. D.
May 2007

Center for Competitive Credit
PERC
POLITICAL & ECONOMIC RESEARCH COUNCIL

Acknowledgements

We are grateful to DataCreìdito for access to anonymized Columbian credit files as well as to TransUnion LLC for their analytical expertise and use of the ACIERTA credit scoring model. In particular, at TransUnion we would like to thank Maria Olga Rehbein, Andres Perez, Abel Puritica, and Marcela Forero. In addition, we would like to thank the dozens of others who contributed useful feedback as earlier versions of this work were presented. In particular, we thank Margaret Miller of the World Bank for her comments.

While the insights offered by these parties were useful, the contents of this report and the opinions expressed therein are solely those of the authors.

TABLE OF CONTENTS

Abstract

Credit bureaus are an institutional solution to the problems stemming from information asymmetries in credit markets. Earlier work established that credit bureaus (especially private ones using positive and negative payment information) lead to greater private sector lending and lower default rates. The studies did not, however, examine the impact of varying rates of participation in reporting that occurs across economies worldwide.

This study assesses the impact of varying participation rates upon access to credit and default rates, with a focus on Latin America. We analyzed the impact of ownership structure (public vs. private), type of credit reporting system (negative-only vs. full-file), and participation in the system (as measured by coverage) on private sector lending as a share of Gross Domestic Product (GDP). The results of this estimation suggest that privately owned, full-file credit bureaus with 100% participation lead to significantly greater lending to the private sector (at least 47.5% greater) than no participation.

The study further demonstrates the importance of participation in a private, full-file credit reporting system through a series of micro-simulations. Using Colombian credit files and a generic scoring model, we simulated the impact of changes in the rate of participation in reporting positive information. Higher participation rates in a private full-file credit reporting system improved the ability of scoring models to distinguish between low and high credit risks, dramatically increased acceptance rates, significantly reduced default rates, and disproportionately increased lending to women and younger borrowers.

Introduction : Theory and Literature

Examinations of credit reporting and financial sector performance have largely focused on the impact of different institutions governing the reporting structure: Can positive information be reported? Does the public or private nature of the credit bureau make a difference? [1] The attention on the institutional features of credit reporting is understandable, especially from a policy point of view. Less attention has been paid to the impact of participation by data furnisher, that is, whether a creditor does in fact report payment information to the bureau.

The provision of payment information, especially positive payment information, is voluntary in almost all the societies that allow it. One consequence of this fact is that participation rates vary considerably across economies. This report examines the impact varying participation rates have upon growth in lending to the private sector with a focus on Latin America, which possesses few barriers to reporting positives and has similar economic institutions but still witnesses considerable variation in participation rates. We approach this issue comparatively and with simulations using Colombian credit files.

Credit bureaus are institutional responses to the problem of information asymmetries, or lack of information, in lending. Economist Ronald Coase suggested long ago that markets will arrive at sub-optimal outcomes, that is, outcomes that do not exploit all trades, if there are costs to transacting.[2] The cost of these transactions include those associated with the resources of searching, contracting, monitoring, and enforcing a market exchange. A large bulk of these costs stem from lack of information and the price of gathering information. Coase's objective was to explain that firms are institutional responses to the costs associated with transacting in the market. The implications have been extended to all kinds of institutions.

Economist George Akerlof examined the consequences of asymmetrical information of goods in markets.[3] When a product can only be considered of average quality because of lack of information to accurately determine

[1] In the context of this paper, the term "positive" refers only to credit data indicative of timely payments. In other contexts, a broader definition has been applied in which "positive" may also include information pertaining to credit limit, outstanding balance, type of credit (such as revolving or installment), date account was opened, and age of debt. This study uses the narrower definition. The term "negative," for purposes of this study, refers to credit payment data that is indicative of late or failed payments, and includes information about delinquencies, collections, bankruptcy, and lien information.

[2] Coase, Ronald H. "The Nature of the Firm." Economica, November 1937, 4, pp. 386-05.

[3] Akerlof, George. 1970. "The Market for Lemons." Quarterly Journal of Economics. 84 (3): 488-500.

its quality, over time products of above average quality will be driven out and their market viability threatened. To demonstrate this, Ackerlof applied the theory of asymmetric information to explain stunted credit markets in India.

The specific dilemma for a lender extending a loan rests in the fact that only a borrower precisely knows his/her intention and capacity to repay a loan. In contrast, the lender must infer the risk profile of the borrower on the basis of far less information. Borrowers have incentives to misrepresent their risk profile, and even when truthful, the lender must still evaluate the claims. The assessment is crucial since a loan involves an agreement to pay in the future.

Joseph Stiglitz and Andrew Weiss (1981) examined the consequences of information asymmetries in lending. They argued that in competitive equilibrium, a loan market may be characterized by credit rationing because of insufficient information. Given the existence of information asymmetries [4] in credit markets, banks must rely on a combination of pricing (interest rates) and rationing [5] to maximize returns, as pricing to cover overall risk results in a dynamic in which markets do not clear. [6] Their argument holds that higher interest rates, while covering some of the risk of borrower default, are also likely to result in adverse selection. That is, higher interest rates attract borrowers seeking to make risky investments with the potential for high rates of return. The price mechanism alone might not clear loan markets because as interest rates increase to compensate for rising risk, riskier applicants are attracted. Similarly, once a loan is made, some borrowers may have incentives not to pay because without information sharing, they can still obtain loans from other lenders. (Collection on loans involves costs, which can vary with the rights of creditors in an economy.) Faced with this moral hazard (the relative lack of penalty for non-payment) and with the problem of adverse selection (higher interest rates attract riskier lenders) that stem from asymmetric information, lenders will ration credit.[7] Jaffee and Russell similarly showed that asymmetric information in lending markets can lead to credit rationing, financial instability, or excessive (non-market clearing) prices depending on the structure of competition.[8]

Specifically, the lending relationships that emerge in response to problems of asymmetric information can help overcome some of these challenges.[9] Lengthy relationships between borrower and lender can provide the lender with some information, albeit limited since it largely covers only the former's experience with the latter. Moreover, relying on such relationships limits access to credit to those already within the system as clients and thereby creates entry barriers to newcomers. Finally, it

[4] That is, borrowers are better aware of their true capacity and willingness to repay than lenders. In the absence of information about the borrower, except what the borrower provides, lenders face the problem of accurately judging the quality or credit worthiness of a borrower when the loan is made and will only discover it over time after credit is extended.

[5] "Credit rationing" refers to the condition in which, among a pool of observationally identical borrowers, some get credit and others do not, leaving the latter worse off than the former.

[6] Joseph Stiglitz and Andrew Weiss. "Credit Rationing in Markets with Imperfect Information," 1981.

[7] Marco Pagano and Tullio Japelli. "Information Sharing in Credit Markets." Journal of Finance. December, 1993: 1693-1718.

[8] Jaffee, Dwight and Thomas Russell, 1976. "Imperfect Information, Uncertainty and Credit Rationing." Quarterly Journal of Economics. 90 (4) 651-666.

[9] For example, see Peterson, Mitchell and Raghuram Rajan, 1994. "The Benefits of Lending Relationships: Evidence form Small Business Data." The Journal of Finance. 49 (1): 3-37.

creates the possibility of monopoly rents as borrowers have fewer outside options than if the information was widely available. (Although it is important to note that information sharing can also act as a complement to relationship lending.)

The macroeconomic consequences of asymmetric information in credit markets, and the behaviors of lenders and borrowers that result from it, are considerable. Stiglitz and Weiss (1992) formally showed that, with credit rationing, monetary policy is likely to have a weak impact in recessionary periods.[10] That is, if banks ration in the face of information asymmetries, an increase in the money supply may only weakly increase available credit in the system. Monetary policy in these circumstances may be far less effective during a recession than in a boom. Furthermore, they showed that the effects of monetary policy vary by sector, according to the extent that the sector is leveraged, such as in construction.

Jaffee and Russell concluded their examination with a suggestion that more attention be paid to the non-price institutions of the loan market "to discover if there may be alternative and better arrangements." [11] Of these, information sharing is the obvious candidate.

Information sharing has been one institutional solution to the problem of asymmetric information and the consequent dilemmas of adverse selection and weak incentives to repay loans. Credit bureaus or registries are the specific institutional mechanism through which information on borrowers is shared by lenders in an economy. Credit bureaus help bridge the knowledge gap between a borrower and a lender by presenting information about a prospective borrower's past credit history, amount of current debt, and other information used to assess credit worthiness, capacity, and risk.

Furthermore, by providing information on delinquencies and defaults that affect a borrower's future ability to access loans, credit registries generate an incentive to pay on time, thereby helping to reduce moral hazard problems.

[10] Stiglitz, Joseph and Andrew Weiss. (1992). Asymmetric Information in Credit Markets and its Implications for Macro-economics, Oxford Economic Papers 44 (4): 694 – 724.

[11] Jaffee, Dwight and Thomas Russell, 1976. "Imperfect Information, Uncertainty and Credit Rationing." p. 665

Empirical studies of credit reporting are relatively recent. These studies have improved our understanding of how credit reports mitigate the inefficiencies in credit markets, and more importantly, how differences in the structure of credit bureaus and credit reporting shape lending.

The earliest econometric work on information sharing found that the presence of credit registries served to increase private sector lending. Pioneering work by Pagano and Japelli showed that private sector lending is greater in countries with credit registries.[12] The study also found that overall risk in countries with credit information sharing was approximately one-third lower than in countries with little or no credit information sharing.[13] In a related examination, Kallber and Udell, using Dunn and Bradstreet information on business credit histories, found that credit registry information was more predictive of small-business loan performance than detailed information in firm financial statements.[14]

Pagano and Japelli considered private and public credit registries to more or less function as substitutes. Subsequent research has shown this is not usually the case. Recent work found significant differences between public credit registries (PCRs) and private credit bureaus and the types of data they collect. Margaret Miller found that the information collected by publicly owned bureaus tends to be less detailed and more oriented towards bank supervision and business financing. By contrast, private credit bureaus are owned and operated within the private sector (frequently some combination of banks), and collect more detailed credit information across an entire range of loans, big and small alike.

A 2002 Inter-American Development Bank/World Bank survey of approximately 200 banks in Bolivia, Brazil, Chile, Colombia, Costa Rica, El Salvador and Peru found that those banks which used private bureau files and primarily lent to consumers or small-to-medium enterprises saw lower rates of non-performance in their

[12] Marco Pagano and Tullio Japelli. “Information Sharing in Credit Markets.”

[13] Marco Pagano and Tullio Japelli. “Information Sharing in Credit Markets.”

[14] Kallberg, Jarl and Gregory Udell, “Private Business Information Exchange in the Unites States.” pp. 203-228 in Margaret Miller ed., Credit Reporting Systems and the International Economy. (Cambridge, MA: MIT Press, 2002)

loan portfolio than banks which did not use bureau data or used public registry data.[15] The same could not be said of those banks that used public credit registry information. A more recent World Bank report confirmed the overall findings of the 2002 IADB/World Bank survey.[16]

Subsequent studies have also evaluated whether or not the inclusion of "positive" data in a credit report has an effect on the distribution and price of credit. Economists John Barron and Michael Staten found that the use of comprehensive credit information—positive and negative credit history—enables lenders to increase lending while better managing their risk. In their simulations, Barron and Staten found that for any given acceptance rate, the use of comprehensive credit information in a generic scoring model yields a portfolio of loans with markedly fewer delinquencies and defaults.[17] By symmetry, for any given default rate, lenders using comprehensive credit reports are able to grant far more loans than lenders restricted to using only negative information when assessing credit risk.[18] These findings have been reproduced by subsequent studies conducted by ACIL Tasman, Margaret Miller, the Inter-American Development Bank, our own studies, as well as those of several others.

While the first generation of empirical economic research on the role of credit information in credit markets provided a compelling case for the important role played by credit bureaus in credit markets (reduced overall risk and promoted growth in private sector lending), second generation empirical economic research has demonstrated that the ownership structure of a credit bureau (public v. private) and the scope of credit data used in lending decisions (comprehensive v. negative only or less robust credit data) are significant variables when considering the growth and health of national consumer credit markets.

[15] IADB, IPES 2005: Unlocking Credit: The Quest for Deep and Stable Bank Lending. (Washington, DC: IADB, 2004) p. 178. http://www.iadb.org/res/ipes/2005/index.cfm.

[16] World Bank, Doing Business in 2004: Understanding Regulation. (Washington, DC: World Bank, 2004) pp. 59-61.

[17] John M. Barron and Michael Staten. "The Value of Comprehensive Credit Reports: Lessons from the U.S. Experience." pp. 273-310 in Margaret M. Miller ed., Credit Reporting Systems and the International Economy. (Cambridge, MA: The MIT Press. 2003) pp. 290-291.

[18] John M. Barron and Michael Staten. "The Value of Comprehensive Credit Reports: Lessons from the U.S. Experience." p. 296.

These treatments have further examined issues confronted by policy makers, for example, the reporting of only delinquencies or the length of time defaults may be kept on file. These formal aspects (such as business practices and regulations) of reporting systems are key to the performance of the financial sector (see below). However, these rules and standard operating practices are only one side of the system.

The other side is *participation* in the reporting system. In most countries, the reporting of elements beyond non-performing loans, usually above specified thresholds, is voluntary. In fact, whether furnishers provide information, whether negative or positive or both, is most often left to their discretion.

The theoretical literature has examined the issue of participation, to some extent, in historical examinations of the evolution of credit reporting systems and norms. One finding is that lenders, for fear of competition or poaching, may underreport or mislead in the information they provide. In response to this, Pagano and Padilla point out that bureaus penalize these lenders by providing inaccurate or incomplete information on their competitors' customers.[19] In other words, bureaus ensure that lenders get from the system exactly what they put into it, and that no firm can game the system to their advantage. They further argue that the norm of "reciprocity" reduces the risk of moral hazard linked to underreporting.

Historically, the provision of information by lenders seems to have developed in tandem with changes in demography and technology. Advances in computing and communication technology have made storage and transmission of *more accurate* payment information in a standardized format less costly. Perhaps more importantly, the weakening ability of banks to access extensive payment information on borrowers resulting from greater labor mobility has created a further incentive to furnish information. These developments have acted to influence decisions by lenders about whether and how much customer payment information to share with credit bureaus.

[19] Marco Pagano and Jorge Padilla. "Endogenous Communication among Lenders and Entrepreneurial Incentives." The Review of Financial Studies 10, No. 1 (Spring, 1997): pgs. 205-236.

It remains the case that lender participation rates in national credit reporting systems vary considerably across economies, both worldwide and throughout Latin America (see below). While the preconditions and dynamics of participation have been theorized, what has seldom been examined in the empirical literature is the impact of varying rates of participation in reporting (and how participation interacts with crucial reporting variables such as the reporting of positives and the ownership structure of the bureau) on financial performance.

In our examination of the Fair Credit Reporting Act in the United States, we simulated the effect of reducing the quantity of information provided to credit bureaus.[20] That study found that as furnishers drop out of the reporting system, the ability of scoring models to differentiate between good and bad risks worsens, with the consequence that the trade-off between market size (acceptance rates) and delinquency rates worsens, and does so in ways that disproportionately impacts the young and minorities.

For rather intuitive reasons, participation is important in credit reporting. But the question of how much participation is important remains unanswered. We examine the issue in the context of credit reporting in Latin America through statistical estimations and through a simulation exercise using Colombian data files in a method similar to that developed by Barron and Staten. We focus on Latin America because its credit registries are extensive and roughly the same age, participation rates vary, and political economies, including reporting regulations, are similar along many salient dimensions. (See the discussion of the methodology below.) Substantively, the lessons are also very important in the context of economic development in emerging markets.

[20] We used a random sample of 3.6 million anonymized credit files and commercial-grade generic scoring models to simulate the impact of lower participation on financial performance. Scenarios A and B in the various simulations in Michael Turner et al., The Fair Credit Reporting Act: Access, Efficiency & Opportunity. (Washington, DC: The National Chamber Foundation, June 2003) passim. Available also online at http://infopolicy.org/pdf/fcra_report.pdf.

The Financial Sector, Economic Development, and Credit Reporting: Latin America in Comparative Perspective

Financial Sector and Economic Development

The importance of a credit reporting system and its qualities to the financial system has been examined extensively in recent years, as noted. Implicit in all of these studies, including this one, is the claim that a well-functioning financial system is crucial for the well-being of an economy. Credit reporting helps the financial system to mobilize savings and allows for the management of risk to facilitate trade and allocate capital, Thereby, helping to foster growth and innovation. Its ability to do so depends on a few factors, but especially on the reduction of information and other transaction costs. Financial intermediaries arise in order to minimize this "friction" in the system.

The research on finance and growth is extensive.[21] Multi-country estimates show that economies with larger financial sectors (under various measurements) have higher rates of growth, greater productivity increases, and faster growing capital stock. The chains are theorized to be direct (allocation of capital to productive investments) and indirect (facilitating exchange, permitting greater corporate control over managers). The consumer credit reporting system is clearly only one part of the system, relating as it does to risk assessment and credit allocation among consumers and small businesses, whose finances are quite often coincidental with the personal finances of their principals. Other factors such as the stock and bond markets are also significant.

Nonetheless, there is ample evidence that private sector lending as a share of GDP impacts overall economic well-being. In cross-country estimations, Ross Levine found that an increase in private sector lending by 30% of GDP can be expected to witness an increase in GDP growth by 1% per annum and increases in productivity and capital stock growth by 0.75% per annum.[22] This is a conservative estimate compared with many others and should be considered in the context of our findings concerning the impact of higher participation rates in private full-file credit bureaus upon growth in private sector lending as a ratio of GDP.

Studies of credit reporting have largely examined the impact of better information sharing on private sector lending. Below, we examine this aspect as well, except we add creditor participation rate as a variable.

The Legal and Regulatory Environment and Credit Reporting

The legal and regulatory environment in which information sharing takes place greatly impacts the structure and development of credit reporting. While, of course, the law could preclude the operation of a credit registry altogether,

[21] For example see Ross Levine, "Financial Development and Economic Growth: Views and Agenda" Journal of Economic Literature, Vol. 25(June 1997), pp. 688–726; Jose De Gregorio and Pablo Guidotti, "Financial Development and Economic Growth." World Development, Vol. 23, No. 3, (March 1995) pp. 433-448.

[22] Ross Levine, "Financial Development and Economic Growth: Views and Agenda." p. 706.

this is rare. The most common manner in which regulations or the law act as an impediment to credit reporting is by either proscribing the reporting of certain types of data or by requiring data to be purged from a consumer's file after a certain period of time. While these rules fall under the rubric of consumer rights, and specifically privacy rights, they often work to the detriment of consumers. Namely, restrictions on the quality and quantity of the data contained in credit reports diminishes the accuracy of the predictions and decisions that lenders make on the basis of that data.[23]

One of our principal concerns is the degree to which credit reporting is "full-file" in Latin America—the degree to which credit reports contain "positive" as well as "negative" data. While many countries around the world prohibit the reporting of "positive" data (for example, Australia), legal proscription is not the reason for the variation among Latin American countries in the amount of positive data provided to public and private credit registries. However, there are laws and court decisions that inhibit the extent to which positive information is shared and recorded.

The fact that there is no strict prohibition against the sharing of positive data is *not* equivalent to the proposition that the legal framework has no impact on the sharing of positives. Moreover, the absence of laws can inhibit data sharing as well, in that there can be—and often is—a reluctance to share information or store information in the absence of well-specified rights, obligations, and recourses of action. Uncertainty about regulation can make economic agents reluctant to take on what amounts to a risk.

In the United States, credit reporting is governed by a comprehensive federal law—the Fair Credit Reporting Act (FCRA). The FCRA addresses both consumer privacy (by restricting the disclosure of data to "permissible purposes") and data accuracy (by allowing consumers to dispute information they believe to be inaccurate and by making furnishers and bureaus accountable for data quality). This approach, characterized by some as a "harms-based" —as opposed to a "rights-based" — approach to data protection, has been largely successful in the U.S. context.[24] The European Union has taken a

[23] Michael A. Turner. Access, Efficiency, and Opportunity. (Washington, DC: The National Chamber Foundation, June 2003)

[24] For discussion of the "harms vs. rights" distinction see Peter P. Swire and Robert E. Litan, None of Your Business: World Data Flows, Electronic Commerce, and the European Privacy Directive (Washington, D.C.: Brookings Institution Press, 1998). For discussion of the economic benefits of the FCRA, see Michael A. Turner. Access, Efficiency, and Opportunity. Information Policy Institute (2003).

somewhat different approach to issues of data protection than the United States. The 1995 EU Data Protection Directive compelled member states to adopt laws that bar the onward transfer of personal data, including the types of information contained in a credit report, without the explicit consent of the subject of that data.

Broadly speaking, data protection laws in Latin America mirror international experience and embody aspects of both the American and European approach. A 2003 study [25] compared the regulatory environment for credit reporting in six Latin American countries—Argentina, Brazil, Chile, Colombia, Mexico, and Peru—with that of Europe and the United States. (These comparisons are inevitable because of the advanced state of data protection law in the EU and U.S. and because of their divergent approaches.) Of these, only Brazil lacked laws specifically bearing on either data protection or the operation of credit bureaus.

The regulatory environment for credit reporting in Latin America (much as elsewhere) is comprised of several tiers: privileges conferred via constitutional right; laws specifically directed at credit reporting; and bank secrecy laws. Constitutional privacy rights do exist throughout Latin America. Colombia and Peru explicitly extend that principle to data. Brazil and Colombia are viewed as having the strongest consumer protection regimes under their respective constitutions. However, while there are substantive protections in many places, there are no laws in Latin America, to our knowledge, that would impact the quality and quantity of data present in private credit bureaus.

Credit Reporting Worldwide and in Latin America

As mentioned, credit registries or bureaus are institutional responses to the problem of asymmetric information in private lending markets. Private credit bureaus first emerged in both the United States and Sweden at the close of the 19th century.[26] Countries such as Austria, Finland, Canada, and Germany soon followed. In Latin America, Brazil, Chile, Peru and Uruguay all established retail payment bureaus during roughly the same period.[27] These early bureaus were typically cooperatives and non-profit ventures set up by local retailers to help determine the creditworthiness of consumers and were also used to assist with debt collection. Notably, retail payment bureaus in Latin America did not contain bank loan information until recently in Brazil.[28] As populations grew more mobile, it became increasingly important for credit bureaus to expand their geographic reach. In the U.S., for instance, by 1906 a trade association was established to facilitate the sharing of consumer data across regions and cities.[29]

[25] Villar, Leon, Hubert. "Regulation of Personal Data Protection and of Credit Reporting Firms: A Comparison of Selected Countries of Latin America, the United States, and the European Union." From Credit Reporting Systems and the International Economy. MIT (2003).

[26] See Marco Pagano and Tullio Jappelli. "Information Sharing, Lending and Defaults: Cross-Country Evidence." Both the United States and Sweden established their first private credit bureaus in 1890. It is possible that informal information sharing mechanisms among lenders and retailers existed prior to this.

[27] The 2005 Report on Economic and Social Progress in Latin America. Chapter 13. "Information Sharing in Financial Markets." Inter-American Development Bank (Washington DC: IADB, 2005) http://www.iadb.org/res/ipes/2005/index.cfm

[28] Robert Hunt "The Development and Regulation of Consumer Credit Reporting in America." Federal Reserve Bank of Philadelphia. (2002) http://www.phil.frb.org/files/wps/2002/wp02-21.pdf

[29] The organization, the Associated Credit Bureaus, Inc., is the antecessor of the Consumer Data Industry Association (CDIA)

Public credit registries (PCRs) were slower to emerge. The Bundesbank established a registry in Germany in 1934 and France established a credit registry by 1946 under the auspices of the Banque de France. Public credit registries are typically operated by a country's central bank, and provision of data is generally a legal obligation.

The primary source of data for PCRs has historically been commercial loans, although in countries where the consumer lending sector is well developed, some consumer payment data may be collected as well. Public credit registries first emerged in Latin America during the 1960s and 1970s in Mexico, Venezuela, and Chile. But more than half of Latin America's public registries only emerged during the 1990s or later, in part due to prior economic instability throughout the region.[30]

Whereas credit reporting is handled exclusively by the private sector in the United States, in Latin America, as in Europe, a variety of arrangements exist. Private credit bureaus of some form operate along side public credit registries in most Latin American countries. Among Latin American nations, only Panama lacks a PCR, whereas only Ecuador and Nicaragua lack private credit bureaus, although there have been efforts to start one.

The form, role, and design of credit registries, whether public or private, naturally reflect the political, economic, regulatory, and technological environment in which they emerge. Nine of the 17 private credit bureaus in Latin America opened after 1989, also owing in part to economic instability in the region.[31] The stabilization of Latin American economies by the close of the 1980s led to growth in the market for medium- and long-term debt. Credit bureau data has little relevance to the business of short-term lending (30-90 days), where information on cash flow and liquidity is far more important than performance on prior loans.[32]

Both private and public credit bureaus have been changing in Latin America. As better storage, reporting and computing technology becomes more widely available, the cost to credit reporting has fallen and continues to fall. Considerable variations do persist, however. Below, Table 1 reports the share of adults with a credit file in both public and private registries. It also reports information on the share of trades that consist of positive payment information. Clearly, the differences are significant.

[30] The 2005 Report on Economic and Social Progress in Latin America. Chapter 13.

[31] Margaret Miller. "Credit Reporting Systems around the Globe: The State of the Art in Public Credit Registries and Private Credit Reporting Firms." From Credit Reporting Systems and the International Economy. MIT Press. 2003.

[32] Margaret Miller. "Credit Reporting Systems around the Globe."

TABLE 1:
Credit reporting coverage and prehensiveness in Latin America

Country	Public registry coverage [33] (% adults with files)	Private bureau coverage (% adults with files)	Positive information on consumer in files (% of total) [34]
Argentina	22.10%	95.00%	25% to 49%
Bolivia	10.30%	24.60%	< 5%
Brazil	9.60%	53.60%	n/a
Chile	45.70%	22.10%	25% to 49%
Colombia	*** [35]	31.70%	75% to 100%
Costa Rica	34.80%	73.40% [36]	< 5%
Dominican Republic	19.20%	34.60%	75% to 100%
Ecuador	13.60%	0.00%	25% to 49%
El Salvador	17.30%	78.70%	10% to 24%
Guatemala	0.00%	9.90%	75% to 100%
Honduras	11.20%	18.70%	75% to 100%
Mexico	0.00%	49.40%	75% to 100%
Nicaragua	8.10%	0.00%	n/a
Panama	0.00%	40.20%	n/a
Paraguay	8.70%	52.20%	n/a
Peru	30.20%	27.80%	50% to 74%
Uruguay	5.50%	80.00%	75% to 100%
Venezuela	16.80%	0.00%	n/a
Mean (excl. absent bureaus)	**18.1%**	**46.13%**	
Max	**45.7%**	**95.0%**	
Min (excl. absent bureaus)	**5.5%**	**9.90%**	

[33] Source: World Bank, Doing Business Database. www.doingbusiness.org/ExploreTopics/GettingCredit/. Information is for 2005.

[34] The data is for 2001, except for Costa Rica, Colombia and Honduras, which is from 2005. From Arturo Galindo and Margaret Miller, "Can Credit Registries Reduce Credit Constraints." March 2001. Research Department. Inter-American Development Bank, Washington, D.C. Additional information from interviews with TransUnion Latin America.
www.iadb.org/res/index.cfm?fuseaction=Publications.View&pub_id=S-143

[35] Colombia possesses a public full-file registry, although the Doing Business database codes it as having zero coverage. Part of the reason for the confusion is that the practices of the bureau make coding it difficult; there is lack of information about the extent to which the bureau shares information with the private bureau and crucially with lenders. Also See Table 3 below.

[36] TransUnion's database contains files on 2.9 million Costa Rican adults (18+ years old). Many of these files have no financial information, but do contain extensive socio-demographic data.

As mentioned, it has been established that positive information matters and that private bureaus make a difference. There is every reason to suspect the differences listed in the table above also make a difference for the lending sector. Simply, they measure the amount of information in a country's financial sector along other dimensions. How much of a difference changes in coverage makes, and by extension participation (see below) makes, especially in the reporting of positives, remains to be tested.

Estimations

One way of assessing the degree to which participation makes a difference is to examine the experience of different economies while taking into account other factors that may shape loan size and performance. Multi-country, quantitative studies are commonly used to examine the impact of information sharing in credit markets.[37]

By and large, these statistical estimations test whether information sharing expands lending to the private sector. Information on consumer loans is not available for many economies, and private sector lending (as measured by a survey of the banking sector [38]) is used as a proxy. Although some studies, based on economies for which consumer loan information is available, have looked at the impact of information sharing on consumer lending as a share of GDP.[39] More recent ones have examined whether information sharing reduces non-performing loans as a share of total loans, using a survey of banks.[40] In order to maximize sample size, we use private sector lending as a share of GDP.

The controls are very important. The ability of creditors to collect on defaulted loans is intuitively crucial in determining whether and how much a bank is willing to lend to a borrower. Previous estimations have considered the impact of legal traditions,[41] wealth, economic growth, the age of the credit registry, the rights of creditors, and, more recently, the impact of ownership structure (or public private differences). Each subsequent look at information sharing adds new variables while keeping, in some form, ones previously established as being important.

This study focuses on how participation in the reporting system affects private sector lending, and how participation interacts with other variables that have been shown to impact borrowing. Our estimates use recent data from the World Bank Doing Business database. The database contains information on *both* public credit bureau coverage and private credit bureau coverage. The database also provides an index on creditor rights and credit information (see below).

37 See Marco Pagano and Tullio Japelli. "Information Sharing in Credit Markets." Also see Simeon Djankov, Caralee McLiesh, Andrei Shleifer, "Private Credit in 129 Countries." NBER Working Paper No. 11078 (January 2005). http://papers.nber.org/papers/w11078.

38 IMF, International Financial Statistics. "Claims on the private sector". Line 52D for 2004.

39 Marco Pagano and Tullio Japelli. "Information Sharing in Credit Markets."

40 Inter-American Development Bank, IPES 2005: Unlocking Credit: The Quest for Deep and Stable Bank Lending. (Washington, DC: IADB, 2005) Chapter 13, p. 178.
www.iadb.org/res/index.cfm?fuseaction=Publications.View&pub_id=B-2005E.

41 Marco Pagano and Tullio Japelli. "Information Sharing in Credit Markets."

We use coverage as a proxy for participation, in so much as more consumers are captured in registries to the extent that more furnishers provide payment information. To see how, consider a list of banks in an economy. Greater participation by the banks in the reporting system results in more coverage because with more participants larger shares of the market are brought into the reporting fold, although it is possible that in some economies coverage could be significantly boosted by the participation of a few furnishers that provide financial services to the vast majority of consumers. Nonetheless, we feel coverage can be a reasonable proxy for participation in the system.

The Doing Business database also provides an index of the legal rights of creditors (on a scale of 1 to 10) based on 10 different variables comprising collateral and bankruptcy law. It measures the extent to which law governing bankruptcy and collateral enable or hinder lending. The incentives to provide a loan clearly depend, in large part, on the ability to recover losses in the event of non-payment. The weaker this ability, the greater are the moral hazard problems in lending.

The Doing Business Legal Rights index comprises three factors concerning rights in bankruptcy and seven factors concerning collateral law. The score is a simple aggregate of the single point assigned for each factor if it obtains, zero if it does not. These factors are: (i) creditors can seize their collateral when a debtor enters reorganization; (ii) creditors are paid first from liquidated assets; (iii) an administrator, rather than management, is responsible for and has effective authority during reorganization; (iv) collateral agreements allow a general description of assets; (v) collateral agreements allow a general description of debt; (vi) security in the property can be taken or granted by any legal or natural person; that is, there is no constraint on the form of the legal person; (vii) there is a unified registry that includes charges over movable property operates;

(viii) secured creditors have priority outside of bankruptcy; (ix) enforcement procedures can be specified in contracts; and (x) out-of-court seizure and sale of collateral by creditors is permitted.[42]

The index also contains an index of credit information based on six variables relating to the breadth and depth of financial data in credit registries. One point is given for each factor that obtains, including: (i) full-file information (both positives and negatives) are distributed; (ii) financial and non-financial credit information (such as from retailers) is available; (iii) more than two years of information is distributed; (iv) reports contain information on loans above 1% of income per capita; (v) borrowers can access their data; and (vi) information on both firms and individuals is available.

[42] From the Doing Business database. http://www.doingbusiness.org/Methodology/GettingCredit.aspx. The index was derived from the methodology developed by Simeon Djankov, Caralee McLiesh, Andrei Shleifer, "Private Credit in 129 Countries." NBER Working Paper No. 11078 (January 2005). http://papers.nber.org/papers/w11078. Our approach is derived from theirs, and our results are broadly consistent with their findings. (See below.)

These sets of aggregated legal and credit information attributes capture many variables considered in previous estimations. The most extensive tests on the impact of the availability of credit information on private sector lending as a share of GDP were conducted by Djankov, McLiesh and Shleifer. Unlike the tests below, they used dummy variables for the presence of a private bureau and for a public bureau. Their creditor rights index had fewer factors, but they also included in their test an inflation variable. They found that the presence of private bureau had a significant and substantial impact on private sector lending, with a resulting difference of 20% to 35% over the period 1978-2003.[43] They further tested the impact of legal origin, whether the legal code was derived from Anglo, Germanic, Scandinavian, French, or Socialist law, and also for contract enforcement days.[44] Some of our estimations also looked at legal origin and it did find a small but measurable impact.[45] However, they did not examine the impact of coverage.

Simple regressions suggest that coverage, and by implication participation, does matter. However, in keeping with the IADB study, it matters to the extent that furnishers *provide information to a private registry*, as shown in Table 2.

There are many reasons why credit bureau ownership structure makes a difference. Public registries were established to assist banking supervisors in assessing the stability of the financial sector. Providing information for lending was a secondary use, albeit one that is quite significant. As noted above, private registries, by contrast, were established precisely to assist lenders in overcoming limited information on borrowers, provide incentives to pay on time, and to generally better assess risk. Toward this end, private firms provide the types of consumer credit

TABLE 2:
Public and Private Bureau Coverage and Private Sector Lending as a Share of GDP

VARIABLE	I	II
Constant	-140.4222 *** (35.0535)	-137.3321*** (34.4511)
Log of GDP per capita (PPP)	17.5727*** (4.4157)	16.9001*** (4.2353)
Legal Rights of Creditors (from 0 to 10)	5.6546*** (2.0737)	5.9317*** (2.0061)
Private Bureau Coverage (0 to 100, as percentage of adults)	0.5540*** (0.1691)	0.5715*** (0.1654)
Public Bureau Coverage (0 to 100, as percentage of adults)	-0.2191 (0.3801)	
R squared	0.6623	0.6604
F-stat (p value)	29.42 (<.0001)	39.54 (<.0001)
Residual Standard Error	30.57	30.4
N	65	65

Errors in parentheses; * p < 0.1; ** p < 0.05; *** p < 0.01

[42] Simeon Djankov, Caralee McLiesh, Andrei Shleifer, "Private Credit in 129 Countries." Table B.

[43] They found that French origin had positive and statistically significant impact in poor countries and a negative but insignificant impact in rich countries. German origin had positive and statistically significant impact in rich countries, as did socialist origins in all countries. Predictably, the longer the contract enforcement days, the lower private sector lending.

[44] It is also likely that the creditor rights variable captures the effect of "legal origin", the national-cultural sources of a country's legal code.

information products that facilitate lending. Moreover, their information is also better formatted and often tailored for risk assessment models. The poor performance of public registries can be understood in this light.

It should be noted that the inclusion of the aggregated "credit information" variable added nothing to the estimation (and similarly it clearly fails to contribute to the estimations below). One chief reason may be that the factors making up "credit information" are attributes that can be found in the practices of private bureaus. (Private bureau coverage and credit information are substantially correlated, 0.568.) The direction of effect is probably complicated, as the ability to gather wider types of credit information allows a private bureau to better perform its main function of serving lenders. The presence of private bureaus encourages the collection of more information and better information practices, as the information is stored for longer periods and comprises wider aspects of the payment universe (such as utilities) in order to serve lenders and not merely regulators interested in reserve requirements.

The estimations in Table 2 do not take into account one factor that has been firmly established as making a considerable difference in lending performance, namely whether the reporting is of only delinquencies and other negatives or whether it also includes positive payment information. To do so, we take into account whether reporting is comprehensive or negative only. For this second set of regressions, we use variables that posit coverage by a combination of private and pubic, full-file and negative only registries. That is, we simply measure the extent of coverage of the credit eligible population by

1. Public negative only files
2. Public comprehensive files
3. Private negative only files
4. Private comprehensive files

The intuition behind testing this constellation of variables is that the *content* of credit reports also must matter for lending. Table 3 shows the results of these regressions.[46]

TABLE 3:
Coverage, Ownership Structure and Comprehensive Reporting (impact on private sector lending as a share of GDP, 2004)

Errors in parentheses; * p < 0.1; ** p < 0.05; *** p < 0.01

VARIABLE	I	II	III[47]	IV
Constant	-142.40*** (35.31)	-139.48*** (35.49)	-133.97*** (35.41)	-130.80*** (32.20)
Log of GDP per capita (adjusted for PPP)	20.31*** (4.65)	18.37*** (4.45)	17.38*** (4.41)	16.85*** (3.87)
Avg. change in GDP (1995-2004)	-1.20* (0.70)	-0.82 (0.64)		
Legal rights of creditors (from 0 to 10)	4.55** (2.07)	4.99** (2.06)	4.68** (2.06)	4.80** (1.97)
Credit information (from 0 to 6)	-3.87 (2.88)			
Private full-file coverage (0 to 100, as percentage of adults)	0.72*** (0.20)	0.60** (0.18)	0.66*** (0.17)	0.67*** (0.16)
Private negative-only coverage (0 to 100, as percentage of adults)	-0.02 (0.86)	-0.13 (0.46)	-0.06 (0.46)	
Public full-file coverage (0 to 100, as percentage of adults)	-0.11 (0.41)	-0.26 (0.40)	-0.17 (0.39)	
Public negative-only coverage (0 to 100, as percentage of adults)	0.16 (0.46)	-0.01 (0.86)	-0.09 (0.86)	
R squared	0.7075	0.698	0.6895	0.6883
F-stat (p value)	16.93 (<.0001)	18.82 (<.0001)	21.46 (<.0001)	44.9 (<.0001)
Residual standard error	29.45	29.65	29.81	29.12
N	65	65	65	65

[46] In the estimations, two outliers that had experienced recent financial crises, Argentina and Uruguay were excluded.

[47]As mentioned above, there is confusion about how to code Colombia's public credit bureau. Regressions assuming a public bureau coverage rate identical to that of the private bureau were also conducted with no real change to the results.

Constant	Log GDP (PPP)	Legal rights	Priv. full file	Priv. neg. only	Pub. full file	Pub. neg. only
-136.01***	17.84***	4.48**	0.65***	0.09	0.33	0.04

As in other studies, wealth and extensive rights for creditors account for a large degree of the variance in lending to the private sector. An extensive basket of creditor rights can contribute significantly to private sector lending, for obvious reasons; lenders are more willing to lend if the chances of recouping the principal is greater in the event of a default. (The expected difference between an economy in which there are none of the rights identified by the World Bank and one in which all 10 rights are present is nearly 45 percentage points.)

However, what is quite telling is the implication that 100% coverage of credit eligible adults by a full-file (or comprehensive) private bureau can be expected to increase private sector lending by more than 60 percentage points of GDP (all else being equal). This figure is substantially larger than that found by Djankov, McLeish, and Shliefer. One likely reason is that they estimated the impact of credit information sharing over a 25-year period, and private sector credit has grown greatly since. In our estimates, removing observations with very high levels of private sector lending, notably the United States and the United Kingdom, resulted in a coefficient of 0.475, which was still significant at the $p < 0.01$ level. (Coefficients on the other variables remained roughly the same.)

It should also be noted that once private bureaus are segmented by whether they have comprehensive (full-file) reporting or negative-only reporting, it becomes *clear that practices also matter*. The value of comprehensive reporting, as established in earlier studies, becomes clear. One factor we did not test is whether the bureau is also owned by it users, namely banks, or specifically the major banks in the country. There is certainly anecdotal evidence that ownership by the large banks may create a problem in the effective use of a bureau (for reasons of stifling competition) and thereby reduce lending. Whether there is more than anecdotal support for the proposition that private bureaus, which are not owned by banks and other users of credit information, are more effective at information sharing, and thereby smarter lending, remains to be tested.

More importantly, these findings are in line with the intuition that more credit information on a larger share of individuals in a society results in more credit being offered them. And the inference that we can reasonably draw is that the more that credit providers and other data furnishers provide information, the more lending we witness in the private sector.

Overall, three properties of a credit reporting system appear to be crucial to the well-being and growth of the financial sector—(i) private ownership, (ii) comprehensive or full-file reporting, and (iii) widespread participation (as implied by coverage). This last factor is at once obvious—since if few participate, and thereby if few consumers are covered, the reporting system will have little effect on the expansion of credit—and crucial, as it points to the fact that legal and regulatory permission to report and keep comprehensive information is only part of what is necessary for an effective reporting system. Given that reporting is voluntary, the actions of potential data furnishers matter considerably.

Simply, the more information made available about both individuals inside and outside the system as a result of greater participation by potential furnishers (which is the only way that more information can become available), the better lenders can assess the risk of loans to a wider range of borrowers. Greater lending is a good thing to the extent that it is a result of ending credit rationing and not merely extending loans to a level beyond borrowers' abilities to afford them. That is, greater information sharing shouldn't lead to *over*-indebtedness.[48]

A second approach to evaluating the impact of greater participation in the credit reporting system examines loan performance and acceptance rates jointly by exploring how more information enables lenders to better distinguish between good risks and bad ones.

[48] Banking regulators in the United Kingdom and Hong Kong have recently suggested that increased information sharing can prevent over-extension and consumer bankruptcy.

Simulations: Methodology

In order to evaluate the extent to which participation by data furnishers in the full-file system matters, we adopted an approach developed in recent years by Michael Staten and John Barron. Their study found that the use of comprehensive credit information—positive and negative credit history—enables lenders to increase lending while better managing their risk. Crucially, their study also established a methodology to test the economic value of using both positive and negative credit information in a credit report—known in the industry as a "full-file" credit report—as opposed to using only negative credit data.[49] This approach has also been used by others.[50]

Barron and Staten constructed research-grade generic scoring models based upon a random sample of anonymized credit reports from Experian. They then redacted data elements from the "full-file" U.S. credit reports to simulate credit data available in the more restrictive Australian "negative-only" credit reporting regime, and re-optimized the model using the modified credit data. Re-optimization is necessary because in a negative-only system, models will be constructed on the basis of negative-only data (the negative-only files could be run through the "full-file" model but this is likely to overstate the impact.)

To simulate the effect of the Australian restrictions, two sets of files are used: a random sample containing all the data and the same sample with the "positive" data struck from the files. Both sets of files are then run through their respective models (Negative-only files are run through the negative-only model and the files containing both positive and negative data are run through the full-file model.) The predictions obtained are then compared against the behavior of consumers over a two-year period. The models are designed to predict the probability that a particular loan will be delinquent 90 days or more over a 24-month period of time. This method allows a test of different reporting systems using actual micro level data with observed performance over time to test predictions.

The advantage of the approach is precisely that it allows many things to be held constant—idiosyncrasies in the law such as the impact of demographic distributions, fiscal and monetary policy, and the business cycle—all factors that can shape access to credit and the performance of the loans. The downside of this 'partial equilibrium' approach is that it does not account for switches in credit decision making, including a greater rationing of credit, the use of greater application data, and other responses to the loss of information. Furthermore, the simulations in some sense presuppose the extension of credit on the basis of what is assumed to be removed, namely positive information. This would be a limit, to be sure, in an instance where the model is not re-optimized. However, we do use a re-optimized model for the negative-only instance. Furthermore, the cross-country evidence above, as well as those found in other studies, provides parameters against which to check findings.

What the simulations allow is a measure of the impact of more robust information on three aspects of the financial system: (i) its efficiency, (ii) its breadth, and (iii) its distributional fairness. It does so, as mentioned, while holding many other factors effectively constant.

[49] Barron, John M. and Michael Staten. "The Value of Comprehensive Credit Reports: Lessons from the U.S. Experience," in Margaret M. Miller ed., Credit Reporting Systems and the International Economy. Cambridge, MA and London, England. The MIT Press. 2003. Pgs. 273-310.

[50] Giovanni Majnoni, Margaret Miller, Nataliya Mylenko and Andrew Powell, "Improving Credit Information, Bank Regulation and Supervision." World Bank Policy Research Working Paper Series, No. 3443 http://www-wds.worldbank.org/servlet/WDSContentServer/WDSP/IB/2004/12/17/000160016_20041217171024/Rendered/PDF/WPS3443.pdf,

The Use of the Colombian Base for Simulations

For our simulations, we used the DataCredito database of Colombian consumer credit reports. The database contains more than 5 million files on Colombia consumers, and covers nearly 32% of the population. The choice of Colombia was intentional. The sole previous modeling of one economy with full-file reporting to simulate another was that of Barron and Staten, who used US credit reports to simulate the Australian economy. (Simulations of the same country under a different reporting system have been conducted for the United States,[51] Brazil, Argentina, and Mexico[52]).

In simulating the reporting systems that are limited in the share of positive information contained in the bureaus, and as found elsewhere in Latin America, the choice of the full-file system used as the base from which the files are drawn is important. Institutional factors, however, also must be taken into account. Thus, the full-file system of the United States may be a good proxy for other Anglo-American societies, such as New Zealand, given elements such as the legal code, creditor rights, etc. However, it may fail to account for salient differences in, for example, Honduras.

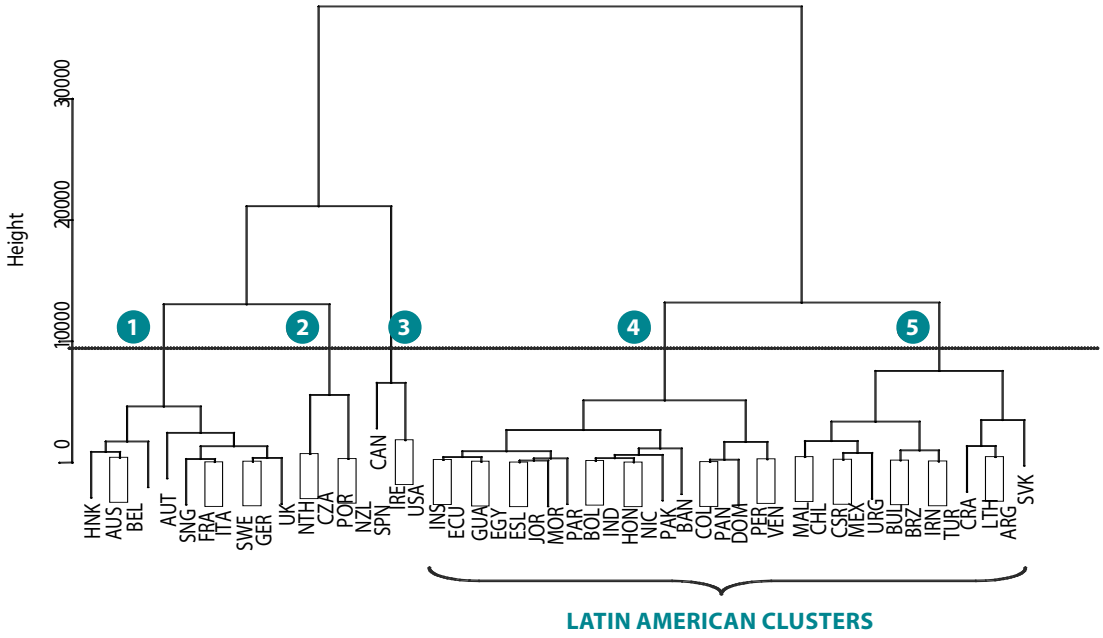
[51] Michael Turner, et al., The Fair Credit Reporting Act.

[52] Giovanni Majnoni, Margaret Miller, Nataliya Mylenko and Andrew Powell, "Improving Credit Information, Bank Regulation and Supervision." World Bank Policy Research Working Paper Series, No. 3443 http://www-wds.worldbank.org/servlet/WDSContentServer/WDSP/IB/2004/12/17/000160016_20041217171024/Rendered/PDF/WPS3443.pdf

A cluster analysis,[53] which tries to order economies in terms of their closeness along four dimensions—per capita GDP (at PPP), rule of law,[54] property rights,[55] and legal origin[56]—shows many of the economies we are attempting to simulate to be "near" Colombia (in cluster "4," Figure 1). The "more distant" Latin American economies (Mexico, Brazil, Argentina, Uruguay, and Chile) are in the adjacent cluster "5". (These clusters also consist of other emerging markets in Asia and Eastern Europe.)

The rationale for these dimensions is that they concern broad institutional features of the economy and the context of economic activity. Legal origins specify a tradition of rights of creditors, as well as of banking regulations. Both property rights and rule of law measure the security in property, which for this exercise also means the ability to collect on loans. This ability shapes lending in so much as it determines incentives to lend. (GDP per capita, of course, measures wealth.)

FIGURE 1:
Economic and Legal Institutions:
Cluster Classifications



[53] We used an agglomerative, hierarchical clustering technique. It used correlations as similarity indicators and using Euclidean distances arranged observations according to those "closest"..

[54] A measure on a scale of 0 to 6 of the law and order tradition in the country based on information from the country-risk rating agency International Country Risk Guide. The measure averages the monthly index of April and October between the years 1982 and 1995. Higher scores indicate a stronger tradition of law and order. Measurements are based on Political Risk Services, International Country Risk Guide (East Syracuse, NY: Political Risk Services, 1996). Aggregated in Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. "Government ownership of banks." Journal of Finance 57, 265–301. (2002) Dataset available at mba.tuck.dartmouth.edu/pages/faculty/rafael.laporta/publications.html

[55] An index of property rights, measured on a scale from 1 to 5, with higher score indicating more protection of private property. "The score is based, broadly, on the degree of legal protection of private property, the extent to which the government protects and enforces laws that protect private property, the probability that the government will expropriate private property, and the country's legal protection to private property." Data from Freedom House. Freedom in the World: The Annual Survey of Political Rights and Civil Liberties 1995-1996 (New Brunswick, NJ: Freedom House, 1996). Aggregated in Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer.. "Government ownership of banks." Dataset available at mba.tuck.dartmouth.edu/pages/faculty/rafael.laporta/publications.htm

[56] Refers to "the legal origin of the company law or commercial code of each country. There are five possible origins: (1) English common law; (2) French commercial code; (3) German commercial code; (4) Scandinavian commercial code; and (5) Socialist/Communist laws." Source: "Foreign Laws: Current Sources of Basic Legislation in Jurisdictions of the World," 1989; and CIA World Factbook 1996." La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer and Robert W. Vishny, "The quality of government, "Journal of Law, Economics and Organization 15, 222-279. (1999) Dataset available at mba.tuck.dartmouth.edu/pages/faculty/rafael.laporta/publications.htm

Colombian Credit Files

Colombian credit files contain name, birthday, and national ID number, with no other socio-demographic information. As mentioned, the files do contain extensive payment and account information, and public record information. Decisions and decision models therefore rely disproportionately on payment and credit history when compared to economies in which demographic information is easily available and legally usable. Unique identifiers are removed to protect the privacy of the data subject.

Colombia credit files also contain extensive *non-financial* account information, especially information on rental payments, utilities and telecommunications. These services are included because, for the most part, they have "credit-like" qualities, in that payment is received often after services have been provided, and the payments occur at regular intervals. And even in the case of rent, non-payment does not immediately result in the loss of services.

One consequence is that delinquency rates in the overall set, which are considerably higher in the non-financial sectors, are higher than what would be expected if examining financial performance solely. Furthermore, the commercial generic scoring model we employed uses both classes of information.

The overall delinquency rate in Colombian consumer lending must be qualified. Problem or delinquent loans accounted for 3.98% of total consumer loans in 2004.[57] A large share of these delinquencies can be attributed to the mortgage sector[58] which has problems collecting largely as a result of considerable legal hurdles in foreclosure.

This information allows us to also examine the impact of changes in data or access to these non-financials, as well as changes in non-performance of accounts in these sectors.

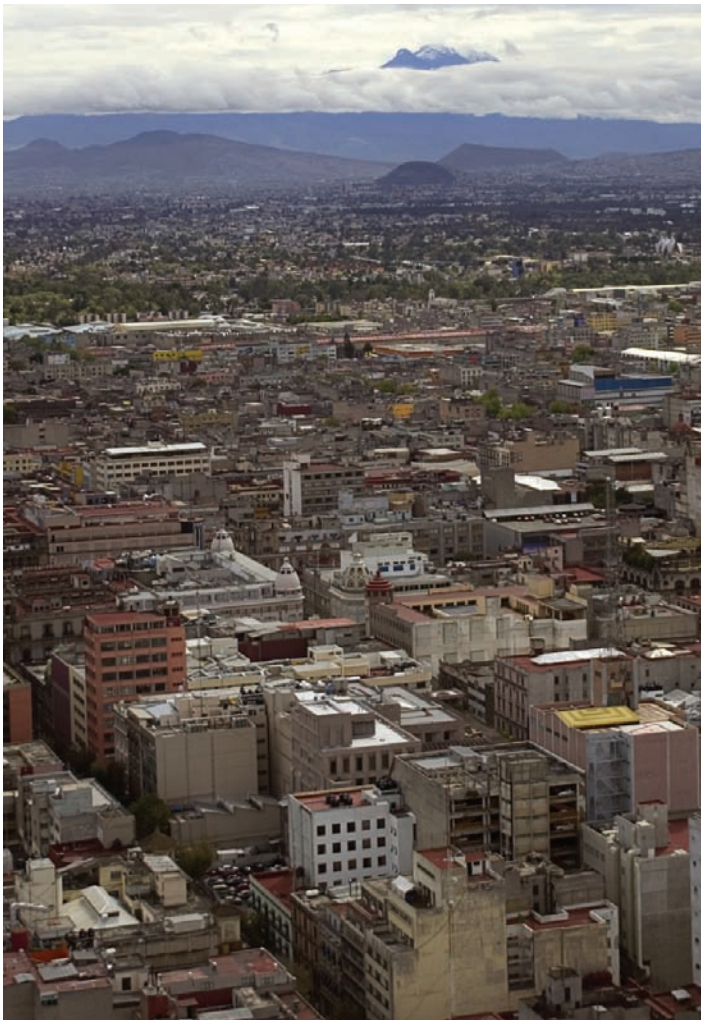
We used DataCredito's entire database of 5.1 million files for the simulations. The files were anonymized and privacy was not violated in the process.. That is, no individual was identified with a file during the scoring exercise or the demographic segmentation of the results, or at any point during this research undertaking.

We constructed four sets of hypothetical files, one for each of the scenarios we considered. For each set, with the exception of the negative-only scenario, we randomly selected the trade lines for which all positive information was purged. An alternative method would have been to choose a data furnisher at random and remove all positives from their lines.

For the simulations, we took credit files from two points in time (July 2004 and July 2005) for each of the data sets. Credit files at July 2004 represent the scored files, the moment when a hypothetical decision is made. The full file at July 2004 and the four hypothetical files for July 2004 (a total of five sets of files constructed from the 5.1 million files in DataCredito's database) were run through (credit scoring) decision models. The timeframe between August 2004 and July 2005 represents the performance period during which we measured the predictive accuracy of the models.

[57] Source: Fitch, Bankscope database.

[58] Mortgages face legal and regulatory impediments, notably a real interest rate ceiling and delays in foreclosure. IMF, "Colombia: Third Review under the Stand-By Arrangement and Request for a Waiver of Nonobservance of Performance Criterion." (Washington, DC: IMF, July 15, 2004) IMF Country Report No. 04/199. p . 9



Scenarios

The scenarios were structured to reflect the level of participation by data furnishers. Above, cross-national evidence indicated that private full-file information leads to wider lending and better loan performance. But we also suggested that the cross-national approach is limited to the extent that there are many aspects of an economy and society that, while having an impact on the credit system, cannot be measured in order to be controlled for—the proverbial "omitted variable bias." Therefore, while these cross-national results give us reason to believe that full-file reporting, especially as performed by private bureaus, improves the development and performance of the financial sector, the results must be supplemented with further approaches. Moreover, the estimations cannot provide the impact of differences in the kinds of participation in full-file reporting, as a result of a lack of data. That is, a system may be full file, but as we have seen in Table 1, there are considerable differences in the share of trades accounted for by positive information. ***This exercise is designed to shed light on the impact of those variations.***

Simulations provide a means of examining the impact of full-file reporting (or different reporting systems) on financial performance through the mechanisms of credit reporting and scoring itself while *de facto* holding all country specific factors constant. That is, it employs the specific mechanism of the way credit is allocated in a society to generate measurable outcomes that can be compared. We can do so while varying the share of positive information in files.

Past simulations have focused on comparisons of full-file vs. negative-only reporting, although, as mentioned, there are exceptions. (Of course, the furthest distance is no reporting at all.) This is because the reporting system, and thereby the contents of credit files, are correctly seen largely as a result of the regulatory system which specifies what types of information can be included. More importantly, the fact that, for example, positive payment information can be legally included does not mean that it will be, especially in systems that are voluntary—as are all those in the Latin American countries examined in this study.

Many data furnishers restrict themselves to providing negative information for reasons of reciprocity and to discipline delinquent borrowers. However, they may be reluctant to furnish positive payment information for fear that their most profitable borrowers will be taken away by competitors, and out of a norm of viewing negative information as necessary and sufficient for an effective credit decision making.

Participation by data providers (financial and non-financial) clearly makes a difference for reasons that are obvious, namely, it determines the quantity of information. Latin American economies witness very different levels of participation in the full-file system, even though there are no real restrictions on sharing positive information. The question is: How much difference is made by participation in a full-file credit reporting system?

To answer this question we developed four scenarios to represent varying levels of participation in the full-file reporting system. Each scenario assumes that different shares of furnishers provide positive information. The four scenarios are:

Scenario 1: 75% of furnishers provide positive and negative information, while the remaining 25% provide only negative information. (25% of randomly selected positive trade lines are removed)

Scenario 2: 50% of furnishers provide positive and negative information, while the remaining 50% provide only negative information. (50% of randomly selected positive trade lines are removed)

Scenario 3: 25% of furnishers provide positive and negative information, while the remaining 75% provide only negative information. (75% of randomly selected positive trade lines are removed)

Scenario 4: all furnishers (100%) provide only negative information

The tiered shares of positive information provision suggest the magnitude of the impact of specific reductions in positive tradelines will have on access to credit for classes of borrowers or default rates given an acceptance rate.

Models

We used three separate scoring models to analyze the effect of changes in the mix of positive and negative data. One is a widely used Colombian commercial grade generic scoring model, ACIERTA, developed by analytic firm TransData S.A. It is, in fact, the sole generic scoring model in Colombia and is noted for being able to well-discern low-risk borrowers from high ones. We also developed a "constrained" version of ACIERTA in order to score the Colombian negative information only files. This version is re-optimized for the negative-only data, unlike the case of the use of ACIERTA in the intervening scenarios, of 75%, 50% and 25% of furnishers providing positive and negative information.

ACIERTA can be used to predict the chances that a consumer will be delinquent on either a specific account or any outstanding account in a one-year period. The model's predictions are measured over a performance period. That is, the model is used to distinguish good credit risks from bad ones. Its predictions, or rather its classifications, can then be compared with actual behavior over a performance period.

In our exercise, the complete set of files and all the hypothetical files except the "negative-only" were scored in July 2004. The scores represent predictions of a consumer's chances of delinquency. We define a delinquency as 90 or more days past due on at least one account in the period between August 2004 and July 2005. How well the model can sort actual good risks from bad ones can be compared across scenarios.

We did not define a delinquency solely on financial accounts, though we did so on non-financials separately. (See below.) This may be a limitation of the study. The reason for defining delinquencies on any account first stemmed from the fact that the model ACIERTA is a generic scoring model used for a wide variety of accounts, financial and non-financial. Moreover, the exercise is designed to ultimately examine a population's access to important services, and in a society, these include utility and telecommunications.

In societies with extensive financial markets, commercial grade financial scores are developed for specific financial products such as mortgages and credit cards. As with some of these, an examination of financial accounts would, in order to show the effectiveness of full-file data, need to restrict or re-weight data to take financial payments into account in different ways when scoring. Second, these non-financial services are "credit like" in that they are promises to pay, and therefore belong in the larger domain of credit worthiness.

As mentioned, we did assess the impact of greater participation in the full-file system on non-financial sectors that participate in the system and crucially use credit information. That is, in a separate set of exercises, we define a delinquency as 90 or more days late on a non-financial account. We included this measure for two purposes, one methodological and the other practical. It is clear that non-financials account for a considerable share of the delinquencies in Colombian files. The non-financial simulation allows us to examine how, for a given acceptance rate, varying levels of reporting positives affects rates of delinquency on non-financial accounts. Focusing on non-financial delinquencies, we hope, will demonstrate to non-financial service providers what greater positive information sharing offers their industry, especially as these services are being used by larger shares of the population in developing societies and promises to help lower income groups enter the formal credit system.

ACIERTA does not provide a decision to "accept" or "reject." Rather, it can be used to rank order consumers/applicants and associate different delinquency rates with these individuals. As such, it allows us to estimate acceptance and delinquency trade-offs under our scenarios. In other words, it shows us how the changes in the provision of positive information alter the choice between extending credit (or even non-credit services as mentioned above) and what share of the loans or services extended will be delinquent. Ultimately, it is left to senior management within each credit granting institution to determine what level of risk or market share to target and pursue.

For these simulations (100%, 75%, 50%, and 25% of furnishers providing positive information), ACIERTA was not re-estimated. That is, it was not re-optimized for the data, although previous comparisons of earlier exercises that re-estimate the model suggest that results would be similar. In our simulation, we did re-estimate in the negative-only simulation.

ACIERTA cannot score files with purely negative information. In order to measure the impact of excluding all positive payment and account information, a "constrained" version of ACIERTA for scoring negative-only was constructed.

The simulations were *not* iterated. Given that the runs were on a very large sample, the deletion of positives on random trade lines was quite likely to have "converged" on accurate values, for all practical purposes.

Evaluation

There are three ways of measuring the value of the different data sets—that is, the value of more information respective to less information.

First, we measure the accuracy of the predictions based on more and less information. The model's score for a credit file is effectively a prediction of the chances that an individual will be 90 or more days delinquent on an account (here, any open account) in the following year. By rank ordering scores, from highest to lowest, we order individuals from those thought least likely to be delinquent in the coming year to those thought most likely to be delinquent in this period.

We can then observe delinquency rates during the performance period (for which we have data of actual behavior). By comparing the delinquency rates for segments of the rank ordering, we evaluate the utility of more data. This comparison can be done in two ways. Assume a lender decides to target an acceptance rate of 20%. We then take the top 20%, as ordered by the model, for each of the scenarios. For each of these sets, we measure the associated delinquency rates, which are then compared. In this manner, we can evaluate the extent to which greater positive payment information helps a lender more accurately predict the risk of lending to a particular borrower.

Practically speaking, these delinquency rates are measures of non-performance, and in the aggregate provide an indication of how healthy the consumer lending segment will be under different reporting regimes, all else being equal. Alhough such an analysis is beyond the scope of this study, such differences are significant for capital adequacy and provisioning requirements as per Basel II.

Conversely, we can assume that a lender opts for a certain performance level, that is, targets an acceptable delinquency rate. In general, the delinquency rate increases as more consumers with lower and lower credit scores are accepted since the ranking is according to model predictions of likelihood of default. For a desired delinquency rate, for example 5%, we measure the associated number of potential borrowers, which is the associated acceptance rate.

By comparing the different acceptance rates under different scenarios of data furnisher participation, we can measure the degree to which more positive payment information about borrowers affects access to credit in the form of an acceptance rate.

For each of these approaches (and each is simply the inverse of the other), we broke down the impact of changes in reporting by age and gender.[59] We evaluated how a reduction in the amount of positive information provided by data furnishers affected the acceptance rate for certain socio-demographic segments, given a target delinquency rate.

The first approach—the default rate for a targeted acceptance rate—measures the relative efficiency of the different scenarios. The second—the acceptance rate for a targeted default rate—measures the breadth of the system or how widely credit is available. The demographic analysis evaluates the relative fairness of the scenarios by measuring how different societal segments are affected by changes in the credit reporting system.

A third approach looks at Type I and Type II error rates associated with each scenario. A Type I error is a false positive, in this instance, a high risk borrower is judged to be low risk. A Type II is a false negative, in this instance, a low risk borrower is judged to be high risk. In the former,

[59] Unfortunately, information on income was unavailable as was any viable proxy for income. Using credit limits on accounts presented a circularity problem, as credit limits were determined in large part by credit worthiness. That is, credit limits as a proxy result in a measurement problem since the ability to assess credit worthiness is what is being tested.

those who do not deserve credit—in the sense that they are risky and the costs of their riskiness will be borne by others—are given credit. In the latter, those who deserve credit—in the sense that they are responsible borrowers—are denied credit.

To measure Type I and Type II errors, we examined the top 25% and bottom 25% of consumers as rank ordered by the models, for each of the four scenarios. The top and bottom quartiles were used because they were proxies for what are to be unambiguous "low risks" and "high risks" in practice. As we approached the median of the distribution, classifying an outcome as an error became harder. For these segments, in each scenario we examined delinquencies and non-delinquencies. This approach helps to measure the relative efficiency and fairness (in a different sense than above) of different reporting regimes. Fairness here refers to granting credit to those who are deserving.

A fourth approach measures "model fit"—the ability of a scoring model to differentiate between good and bad risk borrowers—as gauged by the Kolmogrov-Smirnoff (K-S) statistic associated with each scenario. Unlike the Chi-square or the Gini, the K-S statistic does not assume any particular distribution. The K-S, in this instance, simply measures the maximal distance between the cumulative distributions of bads (or curve of delinquencies) over the score range and goods (or curve of on time payments) across the score range, with a maximum of 100.

For convenience, we have scaled the K-S statistic for the full-file model estimates to 100. The K-S values for the other scenarios measure the relative fit of the model to the data—that is, relative to the full-file model. These provide a measure of the relative "predictive power" of the model for each of the data sets.

Methodological Issues and Limitations of the Analysis

The limits of the approaches have been mentioned above. It should be stressed again that ACIERTA was not re-optimized for the data, except in the negative-only instance. If reporting in Colombia were to move towards less participation in the provision of positive information, new models or re-optimized models would surely result.

The virtue of using ACIERTA is that it is a commercial grade generic scoring model and is superior to research grade models. The estimations for all but the negative-only scenario therefore represent a "worst case" instance.

The two components (commercial grade scoring model and research grade scoring model) are complementary in that they can both point to the direction of change, and when jointly read, offer an accurate sense of the magnitude of the shift. Specifically, the negative only model serves as a check on the other scenarios to the extent that the former serves as a lower bound.

Simulations: Findings

Consequences for "Predictive Power" or Model Fit

As noted above, the Kolmogorov-Smirnoff, or K-S statistic, measures a model's fit with the data, or its ability to tell between two groups (performing loans and 90 or more days delinquent loans, in this case). The K-S makes no assumptions about the distribution of the data. We scaled the K-S on the full-file model (the "100% of furnishers reporting positives" column to 100) and scaled the K-S of the four scenarios to show their predictive power relative to that of the full model. Three of the scenarios (75%, 50% and 25%) were scored on ACIERTA, as was the full model. The negative-only scenario (the column entitled "0%") was scored using a "constrained" version of ACIERTA. For the former three scenerios, the model was not re-optimized for the data. For the negative-only scenario, the model *was* re-optimized. The table also shows breakdowns for demographic segments, specifically age and gender.

The model fit, that is, the predictive effectiveness of the model, degenerates with the loss of positive credit data. Moreover, a model optimized for the derogatory data (the negative-only 0% scenario) also shows a loss of predictive power. With 25% of furnishers reporting only negative information, the model fit declines by nearly 8%. Practically, this means that the model(s) is (are) becoming more and more "incorrect" in their identification of good and bad risks.

TABLE 4:
Scaled k-s, Predictiveness

	Share of furnishers providing positive and negative information (with the remainder providing solely negative information)				
Scenario	***100%***	***75%***	***50%***	***25%***	***0%***[60]
Total					
	100.00	92.42	90.27	87.67	86.78
Gender					
Male	100.00	91.45	89.06	86.51	79.46
Female	100.00	93.06	90.84	88.77	79.03
Age categories					
0-32	100.00	83.45	77.05	72.24	66.90
32-42	100.00	92.41	89.38	86.65	81.18
42-50	100.00	94.80	93.21	91.33	86.13
50-57	100.00	95.74	94.33	92.91	87.23
57+	100.00	94.09	92.97	91.56	87.48

[60] Results of a "constrained" ACIERTA optimized for the data.

When no furnishers are providing positive information, the model's fit declines by more than 13% relative to that of the full-file, ACIERTA instance. The model fit for the in-between scenarios (when 50% report full-file and when 25% report full-file) declines progressively.

Interestingly, the declines in model fit relative to the full-file model for men and women are similar for each scenario. The declines for age segments are another matter. While declines in model fit for those older than 42 is better than the decline in the overall model, and only slightly worse for those 32-42, except in the negative only case, the performance of the model on the files of young people (younger than age 32) is markedly worse. For this segment, the negative-only scenario witnesses a model fit that is only 67% of the 100% reporting model.

While these reductions in KS may not at first seem dramatic, the consequences for default rates and the size and distribution of consumer credit are severe. These impacts are discussed in detail below.

The Consequences for the Cost and Access to Credit

As lenders find it more difficult to accurately discern low and high risks, they will either be forced to accept higher delinquency rates or reduce their acceptance targets. That is, they can opt to reject all but the most clearly credit-worthy applicants, which entails reducing their market size, in order to preserve their margins, or they can accept higher rates of non-performance, requiring them to increase reserves and/or prices of credit. Optionally, they can do both in measure.

Acceptance Rates

Table 5 provides a sense of the magnitudes by which acceptance rates drop for a given default rate across the scenarios, as furnishers provide less and less positive information.

The results of the simulation show that for any given target default rate, especially at lower, healthier levels, as less positive data is reported as acceptance rates drop considerably. For example, if we take the default target to be 5%, roughly the actual share of non-performing consumer

TABLE 5:
Acceptance Rate by Scenario

Share of furnishers providing full-file information (remainder provides negatives only)					
Target default rate	***100%***	***75%***	***50%***	***25%***	***0%***
3%	10.00%	6.64%	4.73%	4.80%	2.56%
5%	41.35%	28.96%	19.28%	9.69%	5.15%
7%	58.82%	45.59%	36.42%	25.71%	13.60%
10%	73.06%	68.09%	68.08%	68.09%	54.97%
12%	77.80%	77.21%	76.49%	75.06%	72.26%

loans in Colombia, the acceptance rate drops from 41.35% to 19.28% when only half of all data furnishers provide just negative information.[61] Even if only 25% of data furnishers provide only negative information, the acceptance rate drops to 28.96%. These drops are significant with small losses of positive information resulting in dramatic reductions in access to credit.

Default Rates

A complementary view of the impact of reduced furnishing is shown in Table 6. It demonstrates what would happen to default rates as data providers report less positive information, for a given acceptance target.

As implied in Table 5, default rates increase for any given acceptance target as furnishers cease providing positive information. From another perspective, default rates *decline* as data furnishers provide more positive information. To get a sense of the default rates, compare the results for the 100% full-file instance and the 50% full-file scenario. As data furnishers cease to provide positive information, the inability of lenders to accurately access risk degrades. In this instance, the default rate increases from 4.89% to 6.67%, an increase of nearly 2 percentage points. Viewed differently, should only 50% of positive payment data be reported, lenders would experience a 30% increase in delinquencies.

TABLE 3:
Default rates by target acceptance
under differing levels of participation

Share of furnishers providing full-file information (remainder provides negatives only)					
Target acceptance rate	***100%***	***75%***	***50%***	***25%***	***0%***
20%	3.52%	3.72%	4.66%	5.91%	8.46%
30%	4.12%	4.62%	5.74%	6.78%	9.06%
40%	4.89%	5.66%	6.67%	7.52%	13.85%
50%	5.86%	6.70%	7.49%	8.22%	14.40%
60%	7.20%	7.73%	8.49%	9.25%	15.30%

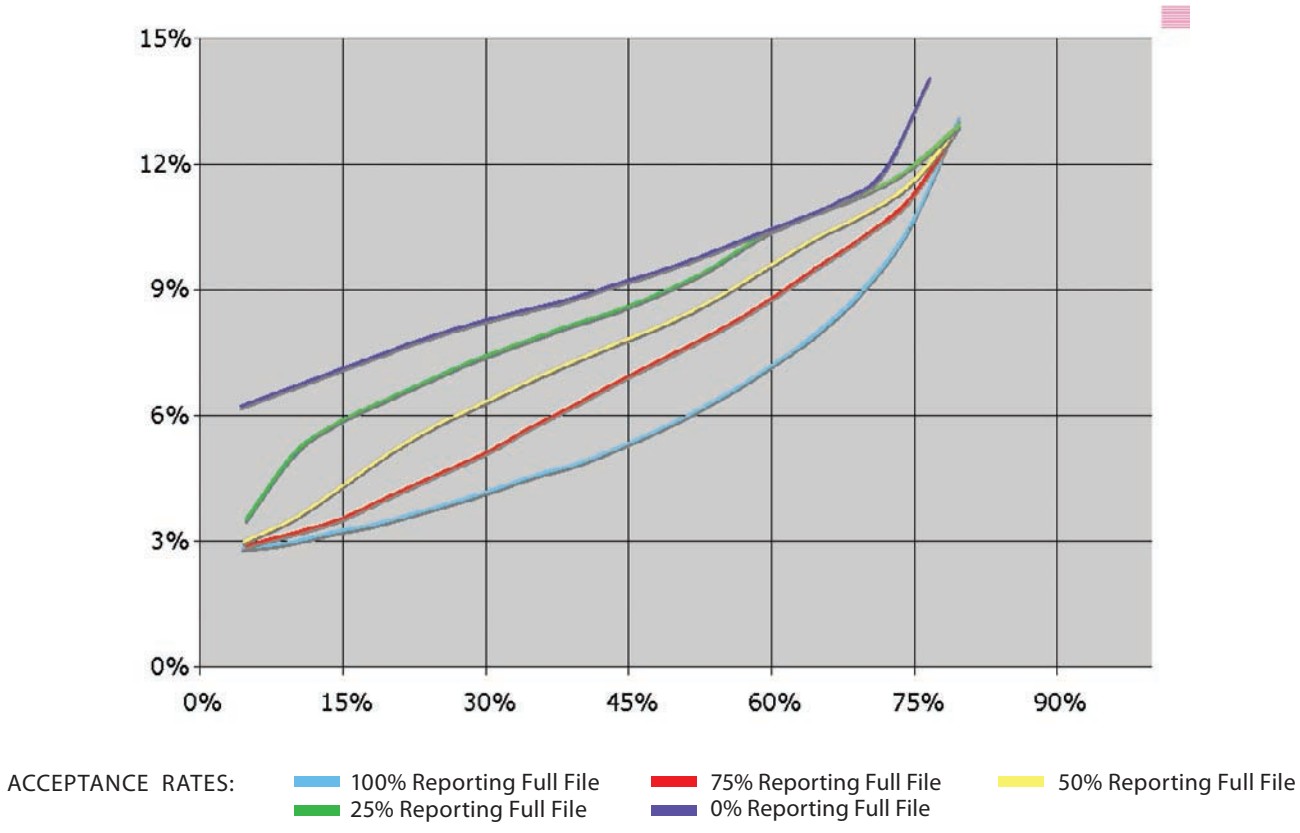
[61] Problem loans as a share of gross loans averaged 4.56 percent between 2000 and 2004. Source: Fitch, Bankscope. Also see IMF, "Colombia: Third Review Under the Stand-By Arrangement and Request for a Waiver of Nonobservance of Performance Criterion." (Washington, DC: IMF, July 15, 2004) IMF Country Report No. 04/199. p. 9

Shifts in the Trade-Off

Tables 5 and 6 depict a shift in the trade-off between acceptance rates and default rates. Figure 2 shows the shift in the curve for the four scenarios and the full-file instance.

As furnishers provide less and less positive information, the "higher" the curve, each acceptance target corresponds to a higher default rate. Furthermore, each default level, in turn, corresponds to a lower acceptance rate. Note also that the differences are significant; the declines in market share are substantial, especially from the full-file instance. At a 6% default rate, the market dramatically contracts (by more than 36%) in moving from a case in which all furnishers provide positive information to one in which 75% do so.

FIGURE 2: Acceptance rate-default rate trade-offs by scenario



Socio-Demographic Distribution of Changes

The change in access to credit is not even across socio-demographic segments. While all segments witness a decline in access to credit, they do so unevenly. Women and young people are more likely to be pushed out of the market by a substantial margin due to a lack of positive information than are men and older consumers with a longer credit history.

Table 7 shows the breakdown of changes in acceptance rates by socio-demographic segments, assuming a default target of 7%.

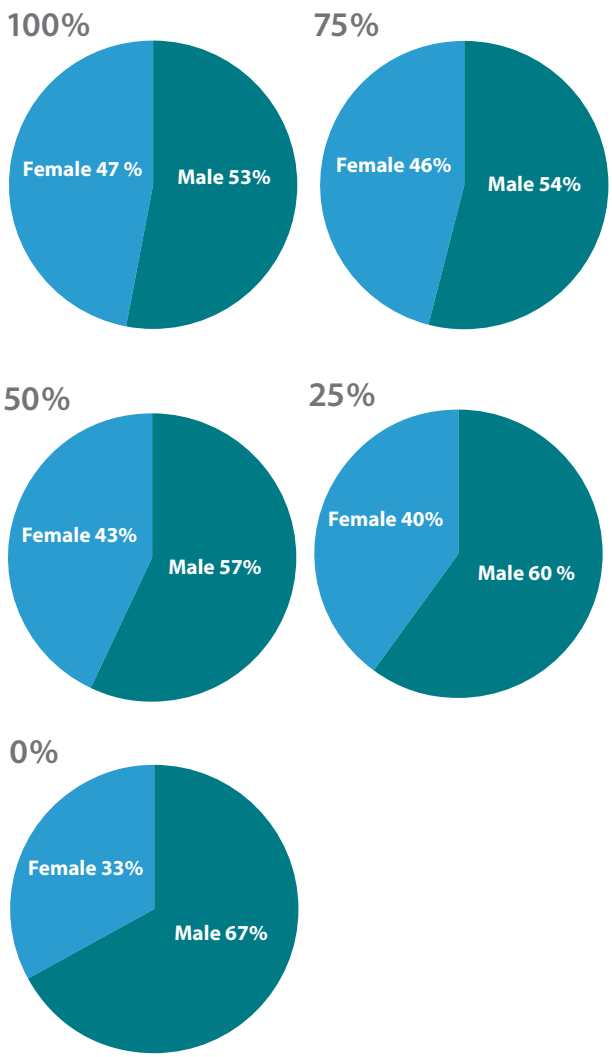
As the results show, women's access to credit falls more precipitously than men's as positive information is removed. In the full-file case, with 100% of furnishers reporting positive information, the acceptance rate for men and women is roughly equal. When only 50% of furnishers provide positive information, the acceptance rate for men falls by 20 percentage points relative to our benchmark, full-file scenario. By contrast, the acceptance rate for women falls by 30 percentage points.

TABLE 7:
Acceptance Rate by Demographic Segment by Scenario (Colombia)

For a 7% default rate	Share of tradelines consisting of both positive and negative information				
	100%	***75%***	***50%***	***25%***	***0%***
Male	64.92%	51.40%	44.31%	33.68%	10.99%
Female	63.20%	42.24%	33.43%	22.30%	5.10%
Age categories					
0-32	16.48%	15.47%	14.20%	8.61%	0.90%
32-42	49.72%	44.75%	28.42%	13.71%	7.67%
42-50	58.31%	45.20%	30.52%	19.14%	12.84%
50-57	62.76%	52.02%	39.61%	19.13%	13.00%
57+	77.13%	72.98%	69.54%	66.49%	20.01%

FIGURE 3: Women as a Share of Borrowers, by Scenario (Colombia)

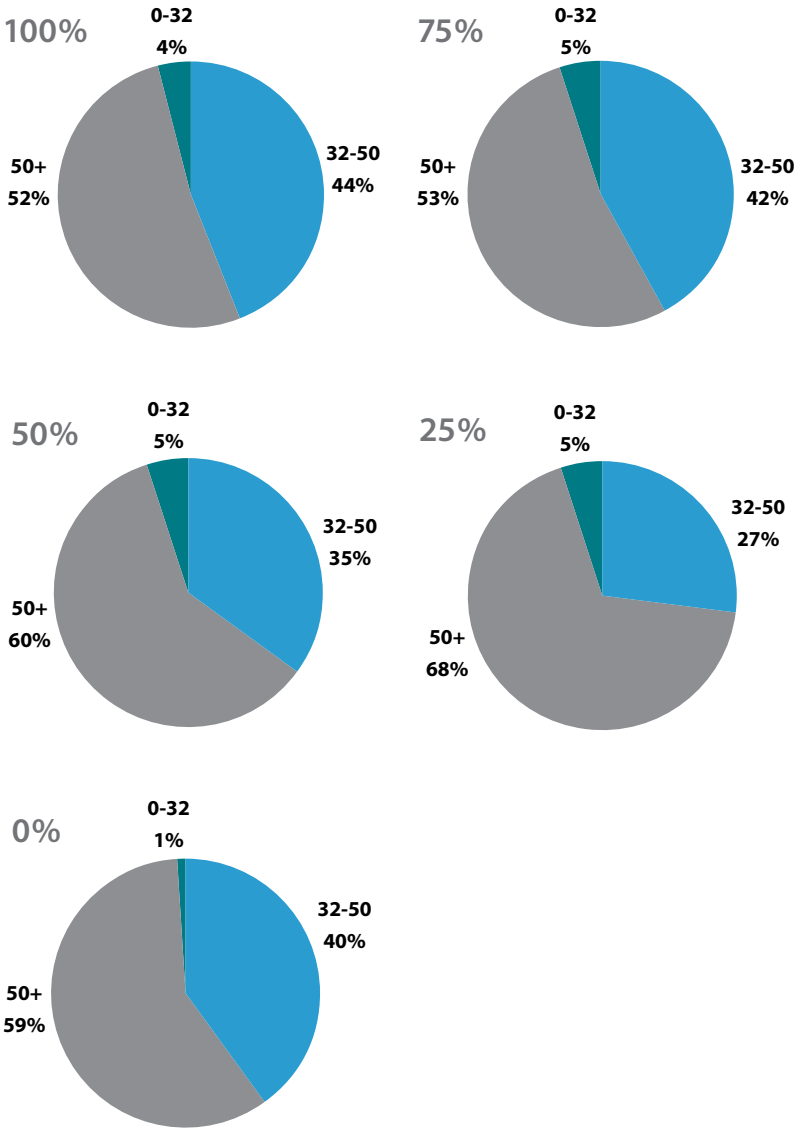
The percentage of women accepted declines steadily as we move across scenarios, and more furnishers cease providing positive payment information. As the share of furnishers providing comprehensive information drops to 50%, the share of women among the accepted (for a 7% default rate) drops from 47% to 43%. When we look at the negative-only, or no positive information scenario, the share of women among the accepted drops to only one-third (33%). Conversely, as more positive information is provided to bureaus, relatively more women are given access to credit.



Why women drop out of the system is not wholly clear. We used an automated scoring system, and thus prejudice is not a factor, especially as gender is not a variable used for decision-making. One possibility is that women, being newer borrowers, and perhaps also more responsible borrowers, are likely to drop out of reports more than men. That is, they may be penalized, as it were, because they are more responsible. Without past information to go on, and without information about existing accounts that would only be registered with delinquencies, women have less access to credit because there is less information about them upon which to make decisions. This finding demonstrates the crucial role played by positive payment information in ensuring fairness in lending. Other dimensions of the relationship between more complete information and fairness in credit granting are discussed below.

FIGURE 4: Age Groups as a Share of Borrowers, by Scenario (Colombia)

The results for changes in acceptance rates by age are also revealing. Young borrowers (less than 32 years old) do not see a steep drop in acceptance rate until 75% of data furnishers cease reporting positive information, at which point their acceptance rate falls from 16.48% in the 100% scenario to 8.61%. Those older than 57 years of age do not see substantial decreases in their acceptance rate until the negative-only scenario. The age groups in between, on the other hand, do see considerable drops in acceptance rates. For instance, in the 50% scenario, the 32-42 age group witnesses a drop of more than 20 percentage points from the full-file case, and the 42-50 age group sees a drop of nearly 28 percentage points. By comparison, the drops in the acceptance rates of those in both the 0-32 and 57+ age groups are modest.



In relative terms, borrowers between 32 and 50 years of age witness steady declines in their share of acceptances throughout, although they witness a recovery in their share in the negative only scenario. (Recall that the negative only is re-optimized for the data.) Again, the reasons are unclear, but most likely the fact of being new to the system leaves *responsible* borrowers with less information to guide decision makers in situations where positive information is not available.

What is clear is that newer borrowers, including disproportionate numbers of women and the young, become disadvantaged.

Error Rates

The loss of the ability to assess risk accurately, which leads to rising default rates and/or worsening acceptance rates, as shown above, stems from the fact that with less information mistakes are more common. The worsening K-S implies as much. Table 8 shows the changes in Type I and Type II error rates for the four scenarios. Here, we see that mistakes, or misjudgments of an individual's risk profile, become more common as furnishers cease reporting positive information. We restrict the results to the ACIERTA-based simulations, that is, all except the negative-only simulation for which a research grade model was constructed.

Table 8 shows that, as positive payment data is lost, mistakes become more frequent.[62] Those who are risky consumers are more likely to get credit, while those who are good risks (not over-indebted and/or have a history of paying responsibly) are less and less likely to be extended credit. The latter group is larger than the former. For the database we used, *approximately an additional 181,000 people who are bad risks would be extended credit as comprehensive credit reporting drops to only 25% of data furnishers from the 100% full-file scenario. Perhaps more importantly, nearly an additional 411,000 people who are good risks and are deserving of credit would be denied access.*

The loss of information results in lower acceptance rates for any given target default rate. However, this result is only part of the picture. Given that false positives increase, the number of those who deserve credit but are denied is even greater than that indicated by simple acceptance rates.

TABLE 8:
Changes in Error Rates (measured as a percent of all credit-eligible adults)

Share of tradelines consisting of both positive and negative information			
	75%	***50%***	***25%***
Type I (false positives, or mistaking a high risk borrower for a low risk one)	+1.00%	+2.22%	+3.31%
Type II (false negatives, or mistaking a low risk borrower for a high risk one)	+3.81%	+5.32%	+7.53%

[62] Given the proprietary nature of the commercial models performance statistics, we are not able to provide the actual rates, only changes.

Non-Financials

It was noted earlier that data furnisher participation in the system not only helps to expand access to financial services and credit, but also to "credit-like" services such as utilities, telecommunications and apartment and home rentals for which payment and receipt of services are not always simultaneous; service providers, therefore, must assess risk. Analysis of this data allows a test of the impact of reporting on these sectors.

This data analysis also allows us to compare the performance of a subset of the data limited by sector with that of whole set. Table 9 shows these results, and indicates that they are largely consistent with the results from Table 2 (although the drop is steeper).

As with financials, more information helps to better predict delinquencies on these services. To the extent that a company must target a default rate, more information allows a firm to provide broader access. To the extent that these services are provided to many consumers via a cross-subsidy, the better identification of low-risk consumers allows for a broader base from which to draw the subsidy pool and thereby lower prices.

Interestingly, Colombia is among just a handful of nations currently using non-financial data in consumer credit reporting. Credit bureaus in the United States have recently been focusing on this matter, and are exploring whether the inclusion of non-financial data in consumer credit reports helps predict the probability of default. This effort is being driven by a desire by mainstream lenders to access data to assist with automated underwriting in new and undeveloped markets, particularly the thin-file and unscorable population segment estimated to be between 35 and 54 million people in the United States. The same type of analysis is ongoing in many emerging markets, characterized by large populations currently outside the mainstream credit system.

The successful use of non-financial data in Colombia in consumer credit reporting, and the lessons learned in this study, are likely to influence credit bureaus in other nations—developed and developing alike—as lenders seek to expand into new markets and improve the performance of their loan portfolios.

TABLE 9:
Non-financial Acceptance Rates, by Scenario (Colombia)

Share of tradelines consisting of both positive and negative information				
Target default rate	***100%***	***75%***	***50%***	***25%***
5%	5.50%	4.00%	2.95%	1.96%
7%	37.30%	29.95%	17.96%	10.07%
10%	61.03%	49.36%	43.14%	36.01%
12%	69.75%	63.27%	57.70%	50.43%

Conclusion

The results described in this study are in keeping with established theory, earlier generation empirical studies, and experience. The provision of more positive payment information helps better identify low credit risks from high ones and reduces misidentifications. The net effects are wider and fairer access to credit for consumers in the form of a larger acceptance rate. More importantly, this wider access can be gained without increases in loan non-performance rates. These simulations show considerable shifts in performance, although, as noted, these changes in performance rates are measured over a series of economic sectors.

At moderate rates of access, such as an acceptance rate of 60%, we found a doubling of the default rate as participation in reporting positives fell to zero. The measured magnitude may be skewed by the inclusion of non-financial sectors, but not necessarily by a significant degree. Simulations by Majnoni, Miller, et al. found that, for the same acceptance target, default rates increased by 28% and by 83% for Argentine and Brazilian files, respectively, as they went from full-file to negative only.[63] Crucially, they restricted their simulation to performance on loans in excess of $20,000USD in the former case and $300,000USD in the latter, loans which are much more likely to be collateralized. These results suggest that the magnitude changes may not be significantly greater with the addition of smaller loans and non-financial lines.

Moreover, we have examined the distribution of credit access as data furnishers withdraw positive information. Disadvantaged groups, which are more likely to be new borrowers, are disproportionately affected. The withdrawal of positive information is likely to leave the files of women and young borrowers "thinner" than those of older males, for historical reasons. Symmetrically, the addition of positive information is likely to assist these groups gain access to credit.

[63] Giovanni Majnoni, Margaret Miller, Nataliya Mylenko and Andrew Powell, "Improving Credit Information, Bank Regulation and Supervision." Table 4, Panel A.

Conclusion

Latin American financial markets have been stagnant in recent years relative to other regions, such as East Asia, and they measure slightly unfavorably to Eastern Europe.[64] Specifically, there are indications that Latin American banks are less efficient than their counterparts elsewhere.[65] This is not necessarily the result of low participation in comprehensive reporting. The more important issue is whether greater participation in comprehensive or full-file reporting can assist the financial sector to expand private sector lending and improve the performance of the retail banking sector.

One specific value of full-file reporting in the region may lie in its potential to limit financial crises, at least to the extent that crises result from adverse selection and moral hazard problems in lending.[66] Latin America has been the most financial crisis prone region in the world for the last 30 years, averaging 1.25 crises per country, with 35% of countries experiencing recurrent crises.[67]

These are benefits which are rarely disputed and have been documented in a number of studies. Yet, many data furnishers—especially in the financial sector—are reluctant to participate in the credit reporting system and provide positive payment information. The great hurdle to wider participation is the fear harbored by lenders that competitors will use their information to poach the more profitable of a lender's customers. In theory, payment information would permit a competitor to identify a profitable client, although, there is a qualification that is needed. The ability of a competing lender to do so depends on whether it can access credit reports for marketing purposes or whether it can only do so when a borrower applies for credit.

Positive payment information permits a competitor to better price a loan when a borrower seeks credit. However, without the ability to access credit files for marketing purposes, a competing lender cannot efficiently go in search of better credit risks or more profitable customers. Even without the ability to use reports for marketing, a lender can offer competitive terms that reflect risk when borrowers seek credit.

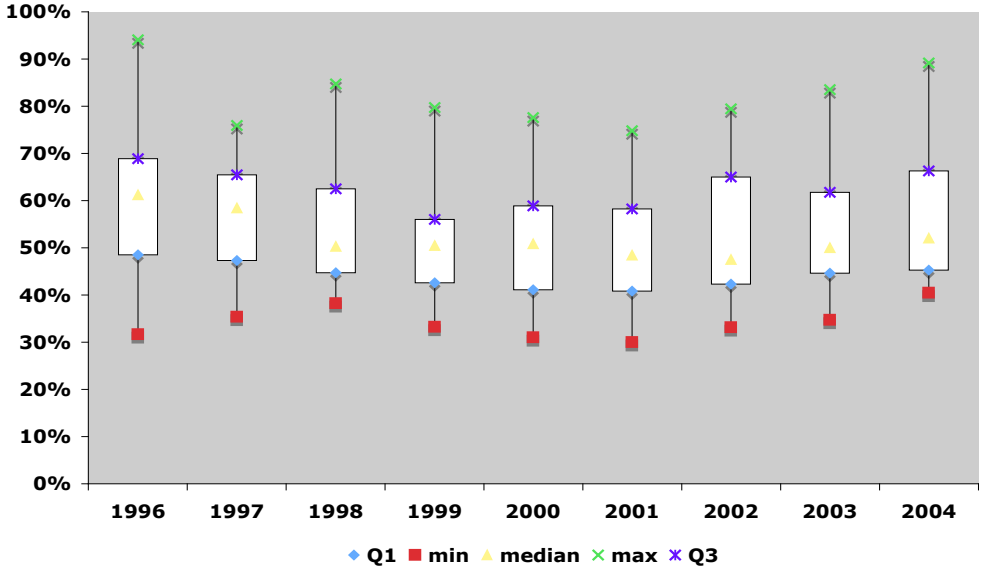
[64] See Alicia Garcна Herrero, Javier Santillбn, et al. "Latin American Financial Development in Perspective" http://ideas.repec.org/p/wpa/wuwpfi/0304008.html.

[65] Alicia Garcна Herrero, Javier Santillбn, et al. "Latin American Financial Development in Perspective." pp. 20-21.

[66] Crises at time result from a shock that exacerbated moral hazard and adverse selection problems. To the extent that these can be reduced by better reporting, crises can be reduced. Frederic S. Mishkin, "Financial Policies and the Prevention of Financial Crises in Emerging Market Countries." In Martin Feldstein, ed., Economic and Financial Crises in Emerging Market Economies. (Chicago, IL: University of Chicago Press, 2003)

[67] The 2005 Report on Economic and Social Progress in Latin America. p. 30, Table 3.1.

FIGURE 5:
Concentration in Banking in 16 Latin American countries. 1996-2004[68]



Competition in an economy's financial sector is the result of many factors: the ease of starting a business, difficulty of capitalization, set up costs and search costs. While the region, like the rest of the world, has seen better consumer financial data systems as a result of new and cheaper technologies for collection, storage, transmission, and computing, and has also experienced increases in the information available on consumers, it has not seen a discernable trend towards less concentration, despite the progress of liberalization. As Figure 5 show, concentration (here measured by a C3, or the market share of the three largest banks) has slightly *increased* since 2000.

Perhaps more important is the question of what is lost as a result of a fear of competition. As the results above suggest, profit rates—in so much as delinquencies and defaults affect profits—are significantly worsened and/or markets are considerably smaller. The overall size of the consumer credit market is also smaller than would be the case if positive payment information were more widely reported. Over time, these losses take their toll on financial markets and loan performance as well as national economic growth.

The case for reporting positive payment data is validated by empirical evidence. In Latin American markets, consumer credit markets would expand without any increase in risk. This could provide a stabilizing buttress against future potential financial crises. In addition, access to credit would become fairer, particularly for responsible borrowers, women, and younger borrowers. Finally, growth in lending to the private sector would lead to growth in Gross Domestic Product (GDP), capital stock, and productivity. In contrast, beyond short-sighted attempts to preserve market share, there is little to back arguments for under-reporting positive payment data.

[68] Source: World Bank, Financial Structure and Economic Development database. www.worldbank.org/research/projects/finstructure/database.htm for 1996-2003; Fitch, Bankscope for 2004 figures.

Annex: the Question of Demographic Information vs. Payment History

With our negative-only Columbian simulation we also sought to compare, albeit imprecisely, a full-file system with little socio-demographic information to a reporting system that consists of negative-only payment data *and* rich socio-demographic information, that is, of the kinds of reporting found in Costa Rica. This raises the question to what extent, if any, can socio-demographic information compensate for the lack of positive payment information? In other words, what is value of socio-demographic data relative to positive transactional data in the context of credit risk assessment?

Towards this end, we also examined 2.95 million Costa Rican credit files. Payment information in these files consists largely of delinquencies of 120 days or greater, with approximately one-third of the trade lines also reporting delinquencies of 90 days or greater.

Costa Rican files have extensive socio-demographic information. The files include information on occupation, employer, length of employment, salary range, parents, martial status, and spouse's name. As with the Colombia files, there is a scoring period July 2004 and a performance period—August 2004 to July 2005. Hypothetical files with socio-demographic information restricted to those found in Colombia file—age and gender—were constructed, as the purpose was (i) to establish a benchmark for comparison with the Colombian negative only simulation, and (ii) to assess how much is gained by the addition of richer socio-demographic information vis-a-vis more extensive payment history data. And as with the Colombian files, unique identifiers were removed to protect the privacy of the subject.

We attempted to answer this question through a structured comparison of Costa Rican files and Colombian files. Both sets of credit files were used to create two hypothetical credit file databases containing common variables; a "Costa Rican restricted" and a Colombian "negative only" (a "constrained" ACIERTA). As mentioned above, basic multivariate, research-grade scoring models were developed for these two data sets. Another model was developed to score the complete Costa Rican files containing all socio-demographic variables. The four results can be compared in three ways: (i) the "Costa Rican restricted" compared to the Colombian negative-only; (ii) the "Costa Rican restricted" compared to the scored Costa Rican complete files; and (iii) the Colombian negative-only compared to the Colombia full-file, ACIERTA instance. In the "restricted" models, we used gender, age, number of days since last 90+ day past due obligations, and number of 120+ day past due obligations. For the complete Costa Rican model, we added average monthly income, marital status and number of dependents.

For the comparison we developed a set of multivariate scoring models using common variables and the full set of available variables for Colombian and Costa Rican files. First, models that relied solely on (i) 90+ day delinquencies, (ii) public records, (iii) gender, (iv) age, and (v) past accounts were developed. (Columbia-restricted and Costa Rica-restricted) Then a multivariate model using all available Costa Rica variables was developed. (Costa Rica-robust)

The comparison was two-tiered. First, the Colombian-restricted and Costa Rican- restricted were compared in order to establish the extent to which the Colombian data can be taken to be a reasonable proxy. Note, this was not a quantitative assessment of comparability but rather a crude one, designed to be suggestive of the degree of comparability. Second, the differences between (i) the Costa Rican-restricted and the Costa Rican-robust and (ii) the ACIERTA constrained and the Colombian ACIERTA full-file were compared. Again, the objective was not a quantitative assessment of the differences but rather a *qualitative* sense of whether full-file payment information was better at predicting risk than rich socio-demographic information. That is, is behavior data more useful than descriptive data?

Were extensive socio-demographic information available in Colombia, a quantitative measure of the value of differing data fields would be possible. However, given the fact that it is not, and further given that this evidence is only a qualitative, albeit strong, indication, the caveats mentioned above should be kept in mind.

There are significant differences in the files and the file sets that make anything more than a suggestive inference unwarranted. First, Costa Rican files exist for those without any financial account, including past ones, as files are created

TABLE 10:
Comparing Lift from Demographics and Positive Payment History, Costa Rica and Colombia

Costa Rica restricted	40.5
Costa Rica complete	49.3
Colombia ACIERTA "constrained"	54.2[71]
Colombia full-file (ACIERTA)	67.3

from identity cards. Without negatives reported, it is difficult to ascertain whether the individual has an account or not; it therefore becomes hard to simply remove those without trade lines. As a result, in the aggregate Costa Rican files understate the rate of delinquencies.

Differences between Costa Rica and Colombia must also be kept in mind, as should the reporting differences mentioned in Table 1. The countries differ significantly in terms of wealth. At purchasing power parity, Costa Ricans are, on average, more than 40 % wealthier than Colombians ($9,886 vs. $6,962, in 2004).[69] Costa Rica and Colombia have averaged similar rates of private sector lending as a share of GDP (27.6% and 25.9% between 1999 and 2004).[70] However, Colombia's trend was downward following the 1999 financial crisis, while Costa Rica's has been growing steadily. Finally, Colombian banks have witnessed better loan performance between 2000 and 2004 than Costa Rica loans, with problem loans averaging 4.56% of total customer loans, compared to 6.99% per total customer loans in Costa Rica. The net upshot of this is that while there are differences, the economies are not so far apart as to be incomparable (see earlier cluster analysis). With that said, there are enough differences to make the results of this simulation purely suggestive.

Any comparison of the magnitude of the changes in performance based on the credit reports here is impossible. The K-S statistics do allow us to compare the value of different data sets. Recall that the K-S is a measure of fit, which for scoring translates into maximizing the separation between the distribution of good risks and bad ones. That is, it is an indicator of an ability to tell low risks from high ones. Table 10 shows the K-S statistic for the 4 simulations.

The relevant measures here are between the differences in the K-S scores for Costa Rica, on the one hand, and Colombia, on the other. The possible implication is that what is gained by much more socio-demographic information is modest when compared to what is gained by much richer full-file payment information. The ability to discern good risks from bad ones (or true positives from false positives) increases considerably in moving from the Colombian negative only to the Colombian full-file scenario (from 54.2 to 67.3). By contrast, socio-demographic information improves the ability to distinguish good from bad risks in Costa Rica files to a much lesser extent.

Again these results are purely suggestive at best, but they are consistent with theory, observation and experience. We can say, however, that the addition of payment history, in the Colombian case, helps to better distinguish true positives from false positives relatively better than the addition of richer demographic information does for the Costa Rican files.

[69] Source: IMF, World Economic Outlook. http://imf.org/external/pubs/ft/weo/2005/02/data/index.htm. Unadjusted for PPP, Costa Rica's GDP was twice as large as Colombia's in 2004, US$4361 compared to US$2149.

[70] Source: IMF, International Financial Statistics. Information from a banking sector survey (52D) was not available.

[71] The K-S statistic may appear high for a negative only instance. High K-S's are achieved through the gains in accuracy of a sophisticated and automated reporting system. Reporting in Colombia is advanced.

DataCrédito
Bogotá, Columbia
Calle 83 No. 14 a 28
www.datacredito.com.co/

TransUnion, LLC
555 W. Adams Street
Chicago, IL 60661
www.transunion.com

PERC
100 Europa Drive, Suite 431
Chapel Hill, North Carolina, 27517
USA
Phone: 919 338 2798
Fax: 212 656 1732
http://www.infopolicy.org