UNIL | Université de Lausanne
Département de finance

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

FINANCE RESEARCH SEMINAR SUPPORTED BY UNIGESTION

"The Information and Agency Effects of Scores: Randomized Evidence from Credit Committees"

Prof. Daniel Paravisini

London School of Economics

Abstract

Information technologies may affect productivity by reducing agents' information processing costs, and by making agents' private information easier to observe by the principal. We distinguish these mechanisms empirically in the context of the randomized adoption of credit scoring in a bank that lends primarily to small businesses. We find that scores increase credit committees' effort and output on difficult-to-evaluate applications. Output also increases in a treatment where committees receive no new information about an applicant, but the score is expected to become available in the future. This effect is uniquely consistent with scores reducing asymmetric information problems inside credit committees. This agency channel explains over 75% of the total output increase. Additional evidence suggests that scores improve productive efficiency by decentralizing decision-making in the organization.

Friday, February 15th, 2013, 10.30-12.00

Room 126, 1st floor of the Extranef building at the University of Lausanne

The Information and Agency Effects of Scores: Randomized Evidence from Credit Committees*

Daniel Paravisini — LSE
Antoinette Schoar — MIT

January 21, 2013

Abstract

Information technologies may affect productivity by reducing agents' information processing costs, and by making agents' private information easier to observe by the principal. We distinguish these mechanisms empirically in the context of the randomized adoption of credit scoring in a bank that lends primarily to small businesses. We find that scores increase credit committees' effort and output on difficult-to-evaluate applications. Output also increases in a treatment where committees receive no new information about an applicant, but the score is expected to become available in the future. This effect is uniquely consistent with scores reducing asymmetric information problems inside credit committees. This agency channel explains over 75% of the total output increase. Additional evidence suggests that scores improve productive efficiency by decentralizing decision-making in the organization.

*We are grateful to Heski Bar-Isaac, Greg Fischer, Luis Garicano, Andrea Prat, Lucia Sanchez, John Van Reenen and the Columbia University, HEC School of Management, Imperial College Business School, European Bank for Reconstruction and Development, Nova School of Business and Economics, Queen Mary School of Economics and Finance, Swiss Finance Institute - USI, University of London - Royal Holloway College, and University of Toulouse seminar participants and the CEPR Workshop on Incentives, Management and Organization, and CEPR/AMID Development Economics Symposium participants for helpful comments and discussions. We wish to thank BancaMia for their support, and Isabela Echeverry and Santiago Reyes for excellent research assistance. We thank IPA for the funding that made this study possible. Please send correspondence to Daniel Paravisini (d.paravisini@lse.ac.uk) and Antoinette Schoar (aschoar@mit.edu).

1 Introduction

The diffusion of information technologies (IT) since the advent of the computer has been associated with increases in productivity in organizations.[1] Empirically identifying the effect of IT on productivity and ascertaining the channel through which it affects performance, however, have proved elusive. The identification problem arises from the fact that IT adoption is usually bundled with other organizational innovations, such as job descriptions, compensation structures or the allocation of authority, which may also affect productivity (see Milgrom and Roberts (1990)).[2] And the main difficulty in isolating the mechanism lies in the dual role played by most IT innovations: they may raise productivity directly by reducing information processing and communication costs, and indirectly through facilitating better monitoring and reducing information asymmetries inside the firm. Distinguishing between the information and agency channels is a key input for understanding the implications of these innovations on firms' internal organization and boundaries.[3] However, existing work attempting to disentangle these two mechanisms has had to rely on ad hoc classifications of whether the information channel or the agency channel will be dominant for each technology.[4]

The present paper uses a randomized control trial design to empirically identify the causal effect of an IT adoption on productivity, and to distinguish the impact through

[1]For early surveys, see Brynjolfsson and Yang (1996) and Brynjolfsson and Hitt (2000).

[2]In addition, IT innovations are typically adopted in response to changes in the environment. In credit markets, for example, it is optimal in theory for lenders to adopt scoring models based on hard information when competition increases (Heider and Inderst (2012)). It is thus difficult to identify the impact of IT innovations from the time series variation of outcomes without imposing structural assumptions (see, for example, Einav, Jenkins and Levin (2012)).

[3]See, for example, Aghion and Tirole (1997) Antras, Garicano and Rossi-Hansberg (2006), and Alonso, Dessein and Matouschek (2008).

[4]For example, Hubbard (2000) identifies two classes of on-board computers in the trucking industry and classifies one as *incentive-enhancing* and the other as *resource-allocation-improving.* Also, Baker and Hubbard (2004) assumes that the introduction of an on-board computer system improves performance through better monitoring. More recently, Bloom, Garicano, Sadun and Van Reenen (2011) classify technologies into communication enhancing and information enhancing, although not for the purpose of separating the information and agency channels.

the information and agency channels. We study how the introduction of a new IT based credit scoring model affects worker effort and output at a bank in Colombia specialized in loans to small enterprises. Prior to the adoption of the scoring model, credit committees evaluate loan applications based on the raw information a loan officer collects in the field from prospective borrowers. The committee determines whether the application should be approved, and conditional on approval, what the terms of the loan should be. If the committee cannot make a decision the loan is passed on to the (higher) manager level for review or an additional round of costly information collection takes place.

We worked with the bank to randomize the roll out of a new credit scoring model in pilot branches. The scoring model provides information about the estimated default probability of a new loan application. Bank headquarters developed this scoring model based on historical loan performance and applicant characteristics such as age, gender, business cash flows and assets, and household expenditure to income. The characteristics used to calculate the score are a subset of those contained in the application file, and are thus observable by the committee before the scores are provided. In the first treatment arm, $T1$, we include the applicant's score in a randomly selected sample of applications submitted to pilot branches before they are evaluated by the credit committee.

There are two ways in which the credit score may lower the cost of acquiring information by the committee. On the one hand it can provide a cheap signal about the default probability that is otherwise costly (or impossible) to obtain through the analysis of the application or through deliberation by committee members. We label this the *information channel*. On the other hand, the credit score may contain information that is privately known by the loan officer —e.g. information obtained from direct contact with the prospective borrower in the field. In this case, the score reduces the asymmetry of information between the officer and other committee members, whose only source of information is the "hard" information contained in the application. Making public the agent's private information reduces the cost of agency problems inside committees that may arise, for example, if loan officers are too conservative, or collude with the borrower. We label this the *agency channel*.

A key innovation in this paper is the design of a second treatment arm ($T2$) aimed at isolating empirically the agency channel. In this randomly selected sample of treatment applications, committees make an interim evaluation of the application *before* observing the value of the score, but knowing that the score will become available to all committee members immediately after the interim decision has been made. This treatment allows us to measure the effect of the expected reduction in information asymmetries between the officer and the committee, while holding constant the information set under which the committee makes decisions. Thus, the effect of $T2$ on interim decisions, before observing the score, measures the agency effect of scores on output.

We find that committees spend more time evaluating applications (more effort) and are more likely to reach a decision on an application (more output) when given access to the credit score in treatment $T1$. The increase in effort is concentrated in marginal —difficult to evaluate– applications that are more likely to be rejected. Despite the upward shift in the difficulty of the tasks performed, the quality of the decisions, measured as the loan approval amounts and the ex post default rate of the loans approved, remains unaltered. The increase in committee output substitutes for other, more expensive, inputs in the loan evaluation process. Namely, when committees reach more decisions they reduce the number of applications that are sent to higher level managers for evaluation.

In the second treatment ($T2$) we find that interim committee output also increases relative to the control group, despite the fact that both have the same information at the time of making a decision. Although output increases even further after the committee observes the score, 75% of the output increase in this treatment group occurs before the score is observed. These estimates imply that the adoption of the scoring model has a first order effect on output through the agency channel.

Looking across both treatments we see that the effect of introducing the scores on the bank's overall output is negligible in the short run —holding constant loan prices, number of employees, pay structure and other features of the bank's organizational design. However, one could conjecture that over time the increase in committee productivity and the time savings at the management level could lead to overall increases in output.[5]

Our results present causal evidence on the role of information technologies in shaping the optimal organization of production. Consistent with the predictions in Garicano (2000), our findings imply that innovations that lower the cost of using information leads to more decentralized decision making inside the firm.[6] By introducing the second treatment arm $T2$ we are able to disentangle the alternate mechanisms that explain information costs inside the firm: technological limitations in information processing and agency conflicts between privately informed workers and the rest of the organization. We show that the agency mechanism explains the bulk of the effect of scores on output, suggesting that IT based monitoring policies may provide an adequate, and cheaper, substitute for policies based purely on monetary incentives.

This is an important empirical conclusion because, in theory, innovations that make agents' information and decisions observable by the principal may have ambiguous effects on the productivity of difficult-to-evaluate workers. In moral hazard contexts where the principal and the agent are symmetrically informed about which actions are appropriate, observing the agents' decisions reduces the cost of inducing effort by the agent (Holmstrom (1979)). In contrast, when agents have career concerns (Holmstrom (1999), Dewatripont, Jewitt and Tirole (1999)), have private information about the productivity of their actions (Prat (2005)), or want to conform to the opinion of the principal (Prendergast (1993)),

[5] See

[6] Our setting is remarkably close to that in Garicano (2000), where workers solve common problems and ask for advice to managers on uncommon ones.

innovations that improve transparency may reduce performance. The results imply that scores help improve the alignment of loan officer incentives with those of the bank and thus suggest that moral hazard costs are of first order importance in our empirical context.

The specific application we focus on, credit scoring, is also of particular importance given the large literature in finance and banking on relationship lending and the role of loan officers in the lending process. This literature has largely focused on the trade off between using soft —less standardized and difficult to communicate— versus hard information. Stein (2002) specifically conjectures that loan officers face weaker incentives in soft information regimes. Our paper provides the first direct evidence to support this conjecture by demonstrating that the adoption of a standardizing technology in the context of a soft information lending process can mitigate agency problems inside the bank.[7]

The rest of the paper proceeds as follows. We provide in Section 2 a description of the tasks and incentives of the credit committees, the characteristics of the credit scoring system. Section 3 describes the experimental design and provides descriptive statistics on the loan applications. Section 4 presents the results of introducing the score on committee output and productivity, Section 5 unpacks the economic mechanism behind the effect. Section 6 concludes.

2 Empirical Setting

The study was implemented with BancaMia, a for profit bank in Colombia that focuses on issuing one- and two-year unsecured loans to micro and small enterprises.[8] The business model of BancaMia, similar to those of other for-profit micro and small business lenders,

[7] A number of empirical studies have analyzed the implications of soft information for bank function and organizational design. See, for example, Berger, Miller, Petersen, Rajan and Stein (2005), and Liberti and Mian (2009).

[8] Although some longer term loans are issued, 90% of the loans have maturities at or below two years.

is based on issuing large numbers of small loans. Only during October 2010, the month prior to the roll out of the study, the bank issued 20,119 loans totalling $US 25,9 million through its 143 branches (average size below $US1,300). Historically the bank relied on a relationship lending model based on two unique sources of information. The first is first-hand verifiable and non-verifiable information about prospective borrowers collected by loan officers in the field. This information collection mechanism is costly but necessary, since micro and small enterprises in Colombia do not have any audited financial statements or other secondary data that a bank could use for credit assessment. The second component is information generated through the repayment history of the riskiest clients: BancaMia offers very small loans to borrowers with high ex ante default probabilities in order to elicit information about their true ex post propensity to repay.[9]

The bank relies on an information technology that allows the loan officer to upload from the field the data collected via PDA (Personal Digital Assistant) devices to a data storage facility in the bank's headquarters. All the information related to an application, including both first hand information collected by the loan officer, past information about the borrower in BancaMia if the borrower has a credit history in the bank at the time of the application, and any external secondary source information (e.g. credit score of the borrower from a private credit rating agency) is put together by the system in a single application file.

2.1 Credit Assessment Process

The application file is then reviewed by a credit committee at the local branch where the loan officer reports to. The committee is composed of the loan officer that collects the information, the branch manager (the loan officer's immediate superior), and one or two additional credit specialists. The credit specialists are typically other loan officers also

[9]For a theoretical analysis and some evidence of the implications of this learning-by-lending approach, see Rajan (1992) and Petersen and Rajan (1995).

employed at the branch.[10] The credit assessment is based on the information that loan officers collected from the borrower in the field and the credit history of the borrower in the bank, if there is one. General information about the industry and a macroeconomic outlook are taken into account as well.

It is important to highlight two unique aspects of the loan prospecting process. First, the loan officer who collects the information makes an active decision to bring an application to the committee, and screens out applicants who do not fit with the desired profile. Thus, applications that reach the committee may not represent the universe of potential borrowers or applications, but only those that have been pre-selected by the field officer.[11] And second, officers are expected to give advice to the prospective borrower on how to fill the application in order to maximize the likelihood of approval. In particular, consistent with BancaMia's learning-by-lending approach, the riskiest borrowers are encouraged to request smaller loan amounts. Thus, requested loan amounts may not represent the borrower's unconditional demand for credit, but rather the demand for credit that maximizes approval probability. A consequence of these risk-adjustments in loan size is that the rejection rate of loans that reach the committee is very low.

Due to this selection in loan prospecting, we must consider the possibility that the introduction of scores changes the officers' incentives to gather data in the field and or changes the applicant selection criteria. For example, if the officers expect scores to reduce the cost of making decisions on marginal cases, they may increase the proportion of marginal applications brought to the committee. In the results section we study how the applicant pool changes during the trial weeks along observable characteristic, such as credit scores and requested loan amounts, and find little evidence of a shift in the applicant

[10] In larger branches there might by a dedicated credit specialist whose only task is to aid in the committee's decision process

[11] An officer may deem an applicant inadequate because it is too risky, or because its financing requirements are too large for BancaMia's business model. All the information regarding potential applicants that do not reach the committee review stage is discarded by BancaMia and is not available for this study.

pool. More importantly, our experimental design, discussed in Section 3, randomizes at the application level once the application reaches the committee to ensure that changes in the application pool do not affect the internal validity of the results.

Once an application reaches the credit review stage, the committee can take four possible actions. First, it can reject the application. Second, it can approve it, in which case the terms of the loan must be decided. The committee may make modifications to the terms of the requested loan —mostly, the amount– in order to improve the acceptance rate. For example, the committee may decide to approve $US500 for a loan application of $US1,000 if the borrower is deemed to be too risky. Since rejection rates are extremely low, the main task of the committee is to decide the "right" loan amount give the expected default probability of the borrower. When a committee takes any of these two actions we consider that the committee has reached a decision regarding an application.

The third action available to the committee is to send the application file to a Regional Manager, whom evaluates the application and reaches a decision.[12] Upper level managers are more skilled and have more experience in credit risk assessment than loan officers or branch managers, and are expected to be more likely to reach the "correct" decision on more difficult applications. The fourth action the committee may take is to send the officer to collect additional information about the borrower. In this case, the committee must take an action on the application in a second round of discussion, after the additional information is collected.

BancaMia managers expressed during informal interviews that the third and fourth actions described above, taken when the committee cannot reach a decision regarding an application, represent a substantial cost to the bank in terms of the opportunity cost of time of managers and officers. It is difficult to quantify these costs precisely. The base fixed

[12]loans above 8 million pesos go directly to the Regional Manager for approval. Randomization insures that this mechanical relationship between loan size and approval level is orthogonal to the scores. Also, adding requested loan amount as a control in the specifications does not change the estimated effect of scores.

wage of a Regional Managers is four to eight times that of a loan officers, which gives a lower bound on the incremental evaluation cost of an application by upper management. Further, the Regional Manager must evaluate the application without the officer that collected the information present and must incur in an additional communication costs to access any soft information not reflected in the application. There are additional delay costs when applications sent up are not reviewed immediately, due to the large volume of applications and time constraints of Regional Managers that supervise between 15 and 80 offices.

Committee member bonus compensation is an increasing function of the number, amount, and value-weighted performance of the loans issued by a branch. Performance pay to loan officers based on lending amount and loan performance is common in most types of lending institutions. The combination of bonuses based on the number of loans and value-weighted loan performance are meant to provide incentives to issue small loans to the riskiest borrowers —compensation based solely on the dollar volume of lending would discourage officers from making small loans.

All the bonuses are calculated on the basis of loans issued, regardless of whether the decision was made by the committee or by the upper level manager. There are two potential reasons for measuring performance based on issued loans only. First, pay based on decisions made would penalize committees for asking questions to the skilled upper level managers, and may lead to too many bad decisions at the committee level. Second, committee members must be compensated for monitoring the performance of the loans after origination, even when the decision to approve is made at an upper level of the hierarchy.

Finally, the interest rates are set centrally based on the type of loan (first-time versus repeat borrower, urban or rural loan). Thus, neither loan officers, committees, or managers have discretion over the price of loans.

2.2 Credit Scores

In 2010, BancaMia developed a credit risk model to establish the statistical relationship between the bank's historic quantitative and qualitative information in loan applications and the repayment performance of issued loans. For the quantitative part of the score, loan officers are asked to collect information such as: gender, age, location, number of years in business, frequency of late payments in past three years (if the loan applicant already has a credit history with BancaMia), level of overall indebtedness, house expenditures as a percentage of total income, among other variables. For the qualitative part, loan officers are asked to collect information based on more subjective variables such as: overall knowledge of business, general sense of the level of organization, quality of information provided, quality of business location, quality of crops being cultivated (agricultural loans only), stability and diversity of income, among other variables.

The stated objective of introducing the credit scoring system was to improve identification of the best and worst clients, decentralize the loan approval process, and reduce the labor costs involved in loan application evaluation. The idea was to add the score as an input to committees' decision process by including the score in the application file.

The score is a proxy for the expected default probability of the loan. Figure 1, panel (a), plots the out-of-sample relationship between scores and default probabilities in the population of loans issued during October 2010 (the scoring model is calibrated using data for loans issued in 2009). For the purposes of the plot, a loan is considered to be in default if interest or principal payments are more than 60 days overdue at six months after the loan is issued. There is a strong positive association between credit scores and and default probabilities, and the tight standard error band implies that scores have a good out-of-sample predictive power for future default. This implies, in turn, that the data collected by loan officers in the field is informative about the repayment prospects of borrowers.

There is a strong negative relationship between default probabilities and requested loan amounts (Figure 1, panel (b)). This relationship is consistent with loan officers screening out large loan applications by risky borrowers, or recommending risky borrowers to request smaller loan amounts. Either way, the observed relationship suggests that loan officers form an accurate prior about the default probability of a borrower before bringing the application to the committee (and before observing the score).

3 Trial Design and Descriptive Statistics

We design a randomized control trial (RCT) with two goals. The first is to measure the causal effect of scores on the effort of the committee, its output and productivity and the total output of the bank. The main hypotheses we want to test is whether and how credit scores lower the cost of decision making for the committee.

The second goal is to decompose the causal effect of scores into two broad mechanisms: information provision and reduction of agency costs. In the pure information mechanism, scores deliver information that the committee members can otherwise only obtain only at a higher cost, e.g. by analyzing the application and deliberation in the committee. For example, the score may contain information that the loan officer already has but that is costly to communicate to others. Alternatively, the score might simplify the processing of information since it automatically assigns weights —based on population data– to the different dimensions of the application which would otherwise require substantial effort by the committee. In our set up, scores do not provide additional information from what is in the application already, since the score only uses inputs from the application form. Instead, scores are an innovation that lower the cost of communicating information to the committee members and of analyzing information in the application file.

In the pure agency mechanism, scores reduce the information asymmetry between the loan officer and the other committee members. In this scenario the score does not provide any new information to the loan officers; it provides the other committee members with information about the repayment probability of a borrower that is privately known by the loan officer. Reducing the information asymmetry inside the committee will affect output if loan officers' incentives are not fully aligned with the those of the bank —for example, if loan officers are more risk loving than the bank, or if it takes a substantial effort to communicate information about a borrower to the committee.

The fundamental distinction between this agency mechanism and the information one is that scores do not bring new information to the committee: all the information is already in the committee but cannot be used effectively due to agency conflicts. We exploit this difference to isolate the effect of scores through the agency mechanism: we design a treatment that reduces the expectation that the informed agent can exploit the information asymmetry, but holds the information set of the committee constant.

3.1 Design

We implemented a pilot program with an RCT design in eight of BancaMia's 24 branches. The branches were chosen to be representative of the average urban branch of the bank.[13] The pilot consisted of randomizing, at the application level, the introduction of scores in the application file at the time of the committee meeting. At the initiation of the discussion of an application in a committee, the research assistants generated a random number to allocate the application into one of three groups, one control and two treatments. Committee members were informed of the group assignment during the evaluation of the application. As argued above, because the randomization took place in real time when the application was being evaluated, we guarantee the internal validity of the results in the presence of potential changes in the application selection criteria or the information gathering process introduced by the pilot program.

[13] BancaMia also operates rural branches, with a larger fraction of loans associated with agricultural micro-enterprises.

In the control group, the committee evaluated the application as before and did not rely on the score. In the first treatment group ($T1$), the committee receives the score before beginning the evaluation of the application. This first treatment allows us to measure the overall effect of scores on committee effort, output and productivity.

In the second treatment group ($T2$), the committee is asked to evaluate the application before receiving the score and chooses an interim action. Committees do not have any additional information when evaluating applications in $T2$ relative to control applications. However, the difference is that the committee members know that the score will become available after choosing an interim action. After recording the interim action, the committee receives the score and can revise it to take a final action (for example, a committee may choose to send an application to the Manager in the interim action, but decide to approve the loan after observing the score). Interim actions in this second treatment allow us to measure how committee behavior changes when the information asymmetry between the committee members is expected to decline, while holding constant the information set available to make the decision.

The main advantage of this design is that it allows measuring the effect of reducing information asymmetry on output regardless of the nature of the underlying agency problem. Moreover, we can make inferences about the nature of the agency problem by looking at whether reducing the information asymmetry reduces or increases the level and quality of committee output. Under relatively weak assumptions, Holmstrom (1979) shows that more information about the agent is strictly beneficial to the principal in moral hazard problems. In contrast, more information about the officer may hurt the bank if it reduces her incentives to work hard to prove her worth (Holmstrom (1999)), or if it increases the officer's incentives to disregard useful private information to act according to what is expected by the bank (Prendergast (1993), Prat (2005)). Thus, the sign of the effect of scores on committee output and default rates allows us to gauge the relative importance of two broad classes of agency problems in organizations.

In a short training workshop before the roll-out of the scores, branch directors and loan officers at the eight pilot bank branches were provided with a general explanation of the credit risk model, the scores, and the objective of the study (researching the usefulness of the score as an input to the credit evaluation process). They were also provided with a detailed description of the three treatment groups and the randomization procedure. We report in Appendix Table A.1 the number of control, treatment $T1$ and treatment $T2$ loans per branch in the study sample.

3.2 Descriptive Statistics

We present descriptive statistics of the applications in the control and treatment groups on Table 1, as well as the p-values of difference of means tests between the three groups. Pre-determined application characteristics —characteristics determined before the randomization takes place– are shown in Panel A. The average requested amount and the applicant score, and the probability that the application is by a first time borrower are not statistically different across the three groups. Figure 2 plots the cumulative distribution of scores and requested loan amounts for the control and treatment ($T1$ and $T2$ grouped together) applications. The requested amount and score distributions are indistinguishable between the treatment and control groups in a two-sample Kolmogorov-Smirnov test for equality of distributions, with corrected p-values of 0.81 and 0.94 respectively. These tests corroborate the internal validity of the randomized design.

Table 1, Panels B through E, presents the statistics for committee and loan outcomes. Some outcomes, such as the time the committee needs to reach a decision, are measured for all applications. Others are measured conditional on a particular action of the committee. For example, the indicator for whether the loan was approved or not is measured conditional on the committee reaching a decision, and the approved loan amount is measured conditionally on the committee approving the loan.

The average time spent evaluating an application in the control group, measured as the difference in the time stamp assigned by the research assistant to the beginning and end of each evaluation, is 4.68 minutes (std. Dev. 3.28).[14] Committees reach a decision (accept or reject a loan) in 89% of the applications, and conditional on reaching a decision, in 0.3% of decisions the committee rejects a loan in the control group. Conditional on loan approval, the committee approves a loan amount different than the requested one in 92% of the applications. The average ratio of approved to requested loan amount is 0.975, but there is substantial variance (Std. Dev. 0.419), and the average absolute value of the difference between the approved and requested amount is $US266, or 17% of the average requested loan amount. The low rejection rate and frequent rate of loan size adjustments are consistent with BancaMia's business model described in the previous section.

Not all approved loans are issued: 75.4% of the loans approved during the pilot program appear as issued in the bank's information system when we collect the ex post outcomes. The bank does not record the reason why the loan is not issued, but presumably the borrower either did not need the loan anymore or got credit from a different source. The default rate among the issued loans —fraction of loans more than 30 days late in repayment measured six months after the loan was issued— is 3.3% in the control group.

Comparing the unconditional outcomes in the treatment and control groups in Table 1, shows that on average committees spend more time reviewing applications in the treatment groups, although the difference is only significant for treatment $T2$ (the difference in average time between $T1$ and $T2$ is not significant). Committees were more likely to reach a decision in both treatment groups than in control applications. None of the

[14]Committees see on average 15 applications per day, which implies that committees spend about 70 minutes per day evaluating applications (net, without considering transitions, breaks, distractions, et cetera).

loan characteristics or outcomes conditional on approval is statistically different in the treatment and control applications.

Table 2 shows the descriptive statistics for applications in the control group conditional on the action taken by the committee —made decision, sent application to the Regional Manager, or sent the officer to collect additional information. On average, the applications for which the committee reaches a decision are for smaller amounts and are more likely to be submitted by first time applicants than applications where the committee does not reach a decision. Applications where the committee reaches a decision are no different in their credit risk (as measured by the score), to those sent up to the Manager, but have a smaller credit risk than those where the officer is sent to collect additional information. Committees spend less time evaluating applications where they reach decisions than when they send the application up to the Manager or the officer out to collect additional information. If one equates evaluation time with effort, this implies that the committee members employ a substantial amount of effort before being able to determine that a decision cannot be reached.

We can also measure final outcomes for applications when the committee did not make a decision during the experiment by tracking the application after the pilot in BancaMia's information system. This allows us, for example, to measure the disbursed amount and the default rate of loans approved by the Manager, or loans approved after a second round of information collection by the loan officer. Note, however, that due to the fact that not all approved loans are issued, we cannot measure the fraction of the applications rejected by the Manager or in a second round by the committee (rejected applications and approved but non-issued applications are confounded in the ex post data). The final loan outcomes differ substantially depending on the action taken by the committee. For example, the default rate is zero for loans sent by the committee to the Regional Manager is and 14.3% for those where the committee sent the loan officer to collect additional information.

The above patterns suggest that there is substantial heterogeneity in the observable characteristics of loan applications, and that this heterogeneity is correlated with the likelihood that the committee can reach a decision. The complexity of the task of reaching a decision and the effort required to reach a "correct" decision is unobservable by the econometrician. In theory, in the presence of task heterogeneity committees should make decisions on the easy-to-evaluate applications and send up to the manager or collect additional information on the difficult-to-evaluate ones (Garicano (2000)). If this is so, the statistics in Table 2 suggest that difficult applications take more time to evaluate, and that applications for larger amounts are more difficult to evaluate. On the other hand, the difficulty of evaluating an application does not appear to be related with the credit risk of the borrower, as measured by the score. We explore further the characteristics of the marginal decisions made by committees in the context of the experimental results.

4 Results

We use the following reduced form equation to estimate the effect of credit scores on committee and loan final outcomes (we delay the discussion on the effect on interim committee decisions for $T2$ until Subsection 5):

$$Y_i = \beta \cdot Treatment_i + \delta \cdot T2_i + X_i' \cdot \eta + \varepsilon_i, \tag{1}$$

where Y_i is an outcome related to loan application i. The variable $Treatment_i$ is an indicator for whether the application is in either treatment group $T1$ or $T2$, and $T2_i$ an indicator equal to one if the application is in treatment $T2$. The vector X_i contains pre-determined application characteristics: applicant's credit score, requested loan amount, a dummy if it is the first loan application of the potential borrower, and the date of the application (in weeks).

4.1 Committee Output and Performance - ATE

For outcomes that are measured unconditionally (e.g. evaluation time or dummy for whether a decision was reached) β measures the Average Treatment Effect (ATE) of having a score as an input to the credit evaluation process, and δ measures the difference in the treatment effects between $T1$ and $T2$.

We present the results of specifications that include predetermined control variables in Table 3 (results without controls are not significantly different, see Appendix Table A.2). The estimated effect of introducing a score on the time it takes to evlaute and application is 0.62 minutes in treatment $T1$, and statistically significant at the 5% confidence level (column 1). This implies that committees spend 13% more time on the average application when scores are available, measured at the mean evaluation time in the control group. Treatment $T2$ has a larger effect on evaluation time than $T1$, but the difference is not statistically significant. This difference is expected since committees must make two decisions in $T2$: an interim one without observing the score and then revise it after observing the score.

This increase in evaluation time comes with more decisions: the proportion of cases in which the committee makes a decision (accepts or rejects an application) increases by 3.3 percentage points, a statistically significant increase at the 10% level (column 2). This implies that when scores are added as an input in the decision process, the number of cases in which committees cannot decide is reduced by close to a third of the baseline proportion of 11% in the control group. Again the difference in the effect between $T1$ and $T2$ is positive but not significant.[15]

One would conjecture these results are the result of committees making more decisions

[15]In Subsection 4.3 we explore further the committee's choice between the four possible actions (approve, reject, send up, or more information) using a multinomial logit.

and taking more time on the marginal cases, i.e. applications that require a higher than average effort to evaluate, when scores are available. To test that this is how the committee is reallocating its time (as opposed to spending more time on all applications), we characterize the effect of scores on the distribution of decision time. Table 3, columns 3 through 5, shows the result of estimating specification (1) using simultaneous quantile regressions for the 25th, 50th, 75th quantiles of evaluation time. The results indicate that only percentiles at or above the median are affected by the introduction of scores, and the point estimates increase monotonically with the quantile. This indicates that scores do not shift the entire distribution of evaluation times. Instead, the availability of credit scores increases the evaluation time on applications that take longer than the median time to evaluate in the first place. This is consistent with scores increasing the time committees spend evaluating more difficult applications.

We next explore whether the effect of providing credit scores varies with ex ante measures of the difficulty of evaluating applications. To obtain objective measures we look at the characteristics of applications where committees did not reach decisions. We saw in Table 2 that committees are less likely to reach decisions for applications for larger amounts, which suggests that application size is correlated with task difficulty. In Figure 3, panel (a) (panel (b)) we plot the relationship between the probability that the committee makes a decision (evaluation time) and the requested loan amount in the treatment and control groups. Treatment appears to increase significantly the probability of reaching a decision and the evaluation time for the largest applications. This confirms that size is related to the difficulty of evaluating applications. More importantly, the coincidence of the effect on the largest amounts across the two outcomes is consistent with the interpretation that committees spend more time on marginal applications.

We repeat in Figure 4 the nonparametric analysis of the treatment effect by scores instead of application amounts. Treatment does not appear to have a heterogeneous impact on applications of different scores. This would suggest that the *level* of the forecast of the default probability is not correlated with the difficulty of evaluating an application, only the *precision* of the forecast is. Put together, the cross-sectional patterns in the treatment effect imply that scores reduce the cost of deciding for any given default probability, and that the reduction is larger for larger loan amounts, where the committee members have more at stake. Although our experimental setting is not designed to establish the link between output and compensation, these results are potentially related to the fact that committee compensation is a (negative) function of the value of defaulted loans, and not their frequency (see Cole, Kanz and Klapper (2012) for randomized evidence on compensation and risk taking by loan officers).[16]

We use the results above to obtain back-of-the-envelope bounds on the additional time it takes to evaluate marginal applications. An upper bound is calculated assuming that the entire increase in evaluation time is due to the applications in which the treatment led the committee to reach a decision when it would not have done so otherwise. This implies that marginal applications require an additional 19 minutes to decide (0.62/0.033). We obtain a lower bound assuming that the increase in evaluation time comes from the 75th percentile of applications, as suggested by the quantile regression results. This implies that marginal cases require an additional 2.5 minutes to decide (0.62/0.25). These estimates imply that deciding a marginal case takes an additional 50% to 400% time to evaluate relative to the average evaluation time of applications where the committee cannot reach a decision in the control group (5.2 minutes).

[16]We also perform a parametric exploration of treatment heterogeneity by augmenting specification (1) with interactions between the treatment dummy and application size and score. These interaction terms are not statistically significant at the standard levels. This is to be expected given the observed patterns in non-parametric plots, where the treatment effect heterogeneity appears to be severely non-linear in loan size, and negligible in score.

4.2 LATE on Conditional Outcomes

For outcomes that are measured conditionally, β in specification (1) represents a Local Average Treatment Effect (LATE) on loans that meet the conditioning criterion. For example, the effect of credit scores on approved loan amounts are measured conditional on the loan being approved. The LATE and the ATE are very likely different in this setting because: 1) the conditioning variable is affected by the treatment status (scores affect the likelihood that the committee makes a decision), and 2) applications where the committee reaches a decision are very different to those when the committee does not (see Table 2). We obtain ATE estimates in the next subsection, when we use secondary data from Bancamia's information system to track the outcomes of loans in the trial without conditioning on the committee's choice.

We present the LATE estimates in Table 4. Conditional on making a decision, the probability that a committee rejects an application increases by 1.1 percentage points in the presence of scores, significant at the 10% level (column 3). This LATE estimate implies an almost fourfold increase in the proportion of applications rejected by the committee relative to the baseline probability of 0.3% in the control group. Given the bank's policy of adjusting the intensive lending margin (loan size) and the documented differences between the marginal and infra-marginal decisions, it is unlikely that this is an estimate of the unconditional effect of scores on the likelihood of rejecting an application. That is, it is unlikely that the score leads the committee to reject applications that it would have otherwise approved. More likely, the effect is concentrated on the marginal applications where the committee made a decision due to the availability of the score (and would have not made a decision otherwise). Assuming that all the additional rejections come from the marginal decisions, the estimate implies that committees reject 24% of the marginal cases they decide on when scores are used as an input $((1.1 - 0.3)/3.3)$. Together with the other findings, the results suggest that more difficult applications are also those that have a higher likelihood of rejection.

Conditional on the committee approving the loan, scores do not have a significant effect on the average approved loan amount, on the probability that the approved loan amount is different from the requested amount, or on the absolute value of the differences between approved and requested amounts (Table 4, columns 2 through 4). So even though committees are deciding on a larger proportion of marginal cases when the score is available, the average credit supply does not change, and neither does the frequency or amount of revisions to the requested amounts. The probability that the loan is issued also does not change with the introduction of scores (Table 4, columns 5) indicating that scores have a negligible effect on the selection that occurs after the approval process is complete. Conditional on the loan being issued, the point estimates indicate that scores increase the non-weighted default probability and lower the default weighted by the amount of the loan, although all the none of the effects is statistically significant (Table 4, columns 6 through 8).[17]

The results so far imply that the introduction of scores in the loan evaluation process increases committee effort, measured as time evaluating applications, and output, measured as final decisions regarding an application. The introduction of scores appears to change the difficulty composition of the problems solved by committees, as it enables committees to reach decisions on applications that are more difficult to evaluate. Despite the upward shift in the difficulty of the tasks performed, the quality of the decisions, measured as the loan approval amounts and the ex post default rate of the loans approved, remains unaltered.

[17]The precision of the estimate on the default probability is low: we would be able to significantly detect changes in the default probability that are larger than 2.5 percentage points.

4.3 Information Collection versus Problem Solving

The data allows identifying two distinct margins through which scores increase committee productivity: 1) by reducing the need to collect additional information from applicants, and 2) by reducing the need to use upper level manager time in evaluating loan applications. We use the following multinomial logistic specification to model committee choice between between approving a loan, rejecting it, collecting additional information, or sending the application to a manager in a higher hierarchical level to make the decision:

$$\ln \frac{P\left(D_i = m\right)}{P\left(D_i = 1\right)} = \beta_m \cdot Score_i + X_i' \cdot \chi_m + \varepsilon_{mi}, \tag{2}$$

where D_i represents the committee choice. We use the committee's decision to approve a loan, $D_i = 1$, as the reference category. All right-hand side variables are as in equation (1). There is one predicted log odds equation for each choice relative to the reference one, e.g. there is a β_m for rejecting a loan, one for collecting more information, and one for sending the application to the manager. A positive estimate for β_m implies that committees are more likely to take action m than to approve a loan in the treatment group relative to the control group.

The results are presented in Table 5. The β_m estimate is positive for the choice of rejecting a loan and negative for, both, the choice to collect more information and to send the decision to the manager.[18] This implies scores reduce both non-decision margins, although the effect is statistically significant only for the choice to send the application to the manager.

To evaluate the economic significance of the effects, we report on the bottom rows of Table 5 the implied marginal effect of treatment on the probability of each choice. Observing a score decreases the probability of sending the decision up to the manager

[18] The coefficients on the treatment regressors β_m are significant at the 1% level in a joint test across the four choices)

by 2.1 percentage points, a 44% decline reduction in the baseline probability that an application is sent to the manager in the control group. Scores reduce the probability of collecting additional information by 1.6 percentage points, a 25% relative to the baseline, but this effect is not precisely estimated.

The results suggest that scores increase committee decision making ability by reducing the degree to which they rely on managers in upper levels of the hierarchy to solve problems. This effect is consistent with models of the optimal organization of production, and in particular, it represents direct evidence that declines in the cost of decision making by workers lead to more decentralized organizations (see Garicano (2000)).

4.4 Overall Performance

So far we have analyzed the effect of scores on committee output, and in this subsection we turn to studying how scores affect total bank output. The effect on total output is ambiguous in the short run because scores and increased committee effort substitute for other inputs to loan production (Manager skills, costly information collection). The short run effects are a partial equilibrium outcome because in the time frame of our study there are endogenous variables that do not adjust to the potential change in the productivity induced by the scores. In particular, the interest rates on loans, which should decline in response to a drop in the cost of lending, are fixed during the study period.

We estimate specification (1) using dependent variables that we can measure ex post from the bank's information system. These outcomes are measured without conditioning on the committee's action, and include applications for which the decision to lend was reached by the Manager after the committee sent the application up, or for which it was reached by the committee in a subsequent meeting (outside of our RCT) after the officer had collected additional information.

The ATE estimate of scores on the probability that the loan is issued is close to zero and not statistically significant (Table 6, column 1). This leads to two observations. First, this result confirms that the observed positive effect of scores on the probability that the committee rejects a loan is the result of committees reaching decisions on marginal loans that would have been rejected anyway (either by the Manager or by the same committee in a second round). And second, since scores do not affect the extensive margin of lending for total output, we can interpret the estimates that condition on the loan being issued (e.g. the effect of scores on loan size) as an ATE.

Scores do not have a statistically significant effect on the measured outputs of the bank (Table 6, columns 2 through 6). The ATE point estimate for the effect of scores on issued loan amount is a precisely estimated zero. The effect of scores on unweighed default is positive and on default weighed by loan amount negative, but both are economically small and estimated with low precision.

These results confirm that the increased output by committees substitutes for other, more expensive, inputs to the production of loans without affecting the overall performance of the decision making process of the bank. This increase in production efficiency is likely to lead to an increase in output once the bank adjusts loan prices and its organizational design (e.g. managers' span of control). Although the partial equilibrium effects on output may be small, we attempt to measure them in the next subsection.

4.5 Loan Prospecting and Branch Output

The introduction of scores may affect the behavior of loan officers in the information collection and loan prospecting stage. For example, in anticipation of the availability of scores in the committee stage of the evaluation process loan officers may change their information gathering effort or shift their attention to particular types of information (from soft information to hard), they may manipulate the entry of data into the system to game the score, or even influence the borrower to change the requested loan amount in the application. In addition, officers may postpone certain types of applications to the committee until the pilot ends.[19]

In this section we investigate the effect the effect of scores on loan prospecting and the selection of applications by looking at whether the experiment changes the pool of applications that reaches the committee. We cannot use the experimental design to study this because the randomization occurs at the committee level, after the selection has occurred. Instead, we perform non-experimental tests that compare outcomes of the pilot branches during the weeks of the trial relative to other weeks, and relative to propensity score-matched non-pilot branches of the bank during the same weeks, using the following specification:

$$Y_i = \gamma \cdot ExperimentWeek_i + Z_i' \cdot \psi + \varepsilon_i, \tag{3}$$

where Y_i is either the score of the borrower, the approved loan amount, or a dummy equal to one if the loan is in default six months after issued. $ExperimentWeek_i$ is a dummy equal to one if the loan was approved during an experimental week in the branch. Z_i is a vector of control variables that includes a full set of branch and week dummies, and branch-specific trends.

We present the results in Table 7, columns 1 through 3, estimated using all the loans approved starting four weeks before the experiment began on the first branch of the pilot (week 41 of 2010), and four weeks after the pilot ended (week 26 of 2011). The propensity score is estimated using the branches' number and total amount of loans approved, average approved loan size and borrower score in October 2010, the month prior to the beginning of the pilot. We find no statistically significant change in the score or requested loan amount of approved loans during experimental weeks. These results imply that the introduction of scores either did not substantially affect the applicant pool, or the change in the applicant

[19]See Heider and Inderst (2012) for a theoretical discussion on loan officer incentives and loan prospecting.

pool was exactly offset by the effect of scores on the composition of approved loans.

We can also use this non-experimental approach to estimate the overall effect of the pilot on total branch output. We estimate specification (3 aggregated at the branch-week level and present the results in Table 7, columns 4 through 9. We estimate specifications using as dependent variable the (log) number of loans issued at the branch, the (log) sum of requested and approved amounts, the fraction of loans that default, and the fraction of debt that defaults (value-weighted defaults). The point estimates on the number and amount of lending are positive, and those on default (unweighed and value-weighted) are negative, but none of the estimates is statistically significant. Overall, we do not find that scores affected total output in the short run. Since scores potentially free up loan officer and Manager time, it is possible that the results are lower bound estimates on the long run effect of scores on total output.

5 The Information and Agency Channels

The results presented so far are obtained from measuring the effect of scores on the final choices by the committee. In this section we turn our attention to evaluating the effect of treatment $T2$ on interim decisions. In treatment $T2$ the committee performs an evaluation of the application and reaches an interim conclusion before observing the score. That is, they chose an action with the same information set as in the control applications, except for the knowledge that the score would become observable by all committee members immediately after choosing.

In theory, we can use this treatment to isolate the agency mechanism. If scores change committee decision making behavior exclusively through the information channel, e.g. by providing information about borrower creditworthiness at a lower cost to the committee, then $T2$ will not lead to an increase in the committees' output (decisions made) before observing the score. On the contrary, in the pure information channel the score and committee effort are complements, so it will be optimal for the committee to put zero effort in evaluating the application before receiving the score, leading to fewer decisions reached in the interim actions. Thus, the pure information channel predicts that the entire increase in committee output relative to the control group will be observed after the score becomes available in treatment $T2$.

In the pure agency mechanism, the future availability of the score reduces the incentives of a privately informed loan officer to distort the loan evaluation process. As a result, the entire effect of $T2$ on output may occur in anticipation of the score becoming available —in the interim action. The direction of the agency mechanism on output is ambiguous a priori, since it depends on the exact nature of the agency problem and how it interacts with the rest of the organization. For example, suppose that loan officers tend hide bad news about prospective borrowers. This could reduce committee output relative to the optimal if committees realize there is missing information and send marginal decisions to the manager. It could also increase output relative to the optimal if the committee wrongly approves more loans than it should. The introduction of scores will increase interim output in the first case, and reduce output in the second.

The information and agency mechanisms are not mutually exclusive, and the results so far based on final outcomes measure the net effect of the two. If both mechanisms are at work, we will observe that $T2$ has an effect on interim actions, and then we will observe committees modify their actions after observing the score.

5.1 Interim Decisions before Observing Scores

We estimate the OLS equation (1) with interim committee decisions as the left-hand side variable, and using for estimation only the control and $T2$ applications. The right hand side variable of interest is a dummy equal to one if application i belongs to treatment $T2$. The coefficient on this dummy measures the effect of making the score available on committee actions *before* the committee observes the score, and thus reflects the gross effect before receiving a new signal about borrower creditworthiness.

We present in Table 8 the results. For comparison, the table includes the estimation of the effect on final outcomes for $T2$, after the committee has observed the score. The effect of the score on the probability of making an interim decision is positive and significantly different from zero at the 5% confidence level (Column 1). The magnitude of the estimated effect is 0.039, smaller than the estimated magnitude on the probability of making final decision, 0.052 (Column 2), but not statistically distinguishable. Committees thus make more decisions in anticipation of receiving the score, and then make even more decisions after observing it. The point estimates suggest that 75% (.039/0.052) of the increase in output occurs before observing the scores.

Conditional on making a decision, committees are also more likely to reject applications during the interim action and before observing the score. In this case, the increase in the probability of rejection in the interim action, 1.5 percentage points (column 3), is larger than the increase in the final outcomes, 1.3 percentage points (column 4), although again the estimates are not statistically distinguishable. Appendix Table A.3 presents in matrix form the transitions between interim and final decisions for all the applications in treatment $T2$, and shows that committees never revise an interim decision to reject an application. This implies that the decline in the point estimate on approval probability between interim and final action occurs due to an increase in number of decisions made and approved. The estimated effect on approved loan amounts and the absolute value of adjustments to requested amounts are negative but not statistically significant.

Finally, we present in Table 9 the multinomial logit model (2) on interim and final outcomes for $T2$. We find that the expectation of receiving a score reduces significantly the probability that committees send an application to the manager in the interim decisions. The effect on the probability declines even more after observing the score, from 2.2 to 3.5

percentage points, although the difference is not statistically significant. In addition, the estimated marginal effects for the probability of collecting more information, although insignificant, suggest that after net of the agency effect, observing the score does not change at all the probability that the committee collects additional information.

Two conclusions can be drawn from these results. First, the bulk of the effect of scores on committee output occurs even holding the information that committees have about the borrower constant. Consistent with the agency mechanism, scores induce committees to make more decisions with information that its members already possess, and this information is relevant for deciding and rejecting marginal applications. These findings are consistent with prior evidence in other settings that shows that privately informed loan officers tend to hide bad news about borrowers (see Hertzberg, Liberti and Paravisini (2010)).

Second, scores also affect output through the pure information channel, but its magnitude is small relative to the agency channel, and the information in the score is not useful for rejecting marginal applications. The lack of an effect on the probability of rejection may be justified in this environment because of the emphasis of the bank in adjusting loan sizes to the perceived borrower risk. However, scores do not appear to have an effect on average loan size or adjustments thereof.

5.2 Information Effect: Intensive Margin

We can also explore further the pure information effect on the intensive margin of lending by comparing the loan amounts approved in the interim and final decision for the applications in treatment $T2$. The scatterplot of the two decisions, shown in Figure 5, indicates that although most of the amounts remain the same after observing the score (77.2%), in 16.2% of the cases the approved amount is revised downwards, and in 6.6% revised upwards, after observing the score. This plot indicates that scores have a substantial influence in the size of loans approved, but the effect is mean zero and thus not captured by average treatment effect estimates.

A mean zero effect on lending may improve the quality of lending if loans below (above) the 45 degree line in Figure 5 are more (less) likely to default. That is, if committees reduce the loan size for loans that default ex post and viceversa. In fact, the opposite is true. The fraction of loans that default that are below, at, and above the 45 degree line in Figure 5 are 0%, 3.8% and 12% respectively. This implies that committees increased the size of loans that ended up defaulting and decreased the size of loans that did not. If default probabilities are not affected by the loan amount then the loan size adjustments done after the committees observed the scores in treatment $T2$ are value destroying. If default probabilities increase with loan size, then it is possible that the changes are value enhancing but it is difficult to evaluate without a proper counterfactual.

The bottom line of this final discussion is that we find little evidence to suggest that the pure information mechanism has first order implications for loan outcomes. This implies that most of the relevant information contained in the scores is already known by the committee members, and that the fundamental problem of the bank is to provide incentives so that the information is used effectively. The results suggest that innovations that reduce informational asymmetries inside the committees may be an efficient way of providing such incentives.

6 Conclusions

In this paper we use a randomized controlled trial to identify the incentive effect of an information technology innovation at a Colombian bank that specializes in lending to small enterprises. We measure the effect of providing credit scores on the productivity of credit evaluation committees. We find that credit scores increase the effort committees put into solving more difficult problems. As a result, scores increase committee's overall output and reduce the need for higher-level manager involvement in the decision process. Thus, the paper presents direct evidence on how information technologies can lead to the decentralization of decision processes inside organizations.

There are two potential mechanisms that drive the increase in committee productivity: (1) reducing committees' information processing costs (information channel), and (2) making loan officers' private information easier to observe by the committee members (agency channel). To disentangle these two channels in treatment $T2$ we do not provide the committee with a score, but we tell them that the score will be made available to all committee members soon after they have reached a decision. We find that the expected future availability of the score increases committee output. Moreover, 75% of the total increase in output occurs before committees see the score in this treatment. This suggests that scores increase output by reducing asymmetric information problems between the loan officer and the committee.

These findings have interesting implications regarding the design of incentives inside organizations. IT based solutions that increase the ease with which the principal can monitor the actions of the agents may have first order effects on productivity and organizational design. These results are particularly surprising in our context, since even without the score the supervisors would have been able to ultimately observe loan officers' choices, for example, when they review the loan officers' performance and bonus payments on a quarterly basis. So the intervention improved the *immediacy and ease with which the principal can monitor the agents but not* whether they get reviewed. It also affected how salient the information is to both the agent and the principal, and thus related to the work by Cadenas and Schoar (2011) who change the frequency of incentives to help loan officers overcome procrastination issues. It is suggestive that these relatively subtle changes in how agents are monitored provide very such significant changes in behavior. As such IT solutions may represent an effective and low cost alternative to steepening or increasing monetary incentives.

References

Aghion, Philippe, and Jean Tirole (1997) 'Formal and Real Authority in Organizations.' *Journal of Political Economy* 105(1), 1–29

Alonso, Ricardo, Wouter Dessein, and Niko Matouschek (2008) 'When Does Coordination Require Centralization?' *The American Economic Review* 98(1), 145–179

Antras, Pol, Luis Garicano, and Esteban Rossi-Hansberg (2006) 'Organizing Offshoring: Middle Managers and Communication Costs.' *NBER Working Paper.* 12196

Baker, George, and Thomas Hubbard (2004) 'Contractibility and Asset Ownership: On-Board Computers and Governance in U.S. Trucking.' *Quarterly Journal of Economics* 119(4), 1443–1479

Berger, Allen, Nathan Miller, Mitchell Petersen, Raghuram Rajan, and Jeremy Stein (2005) 'Does function follow organizational form? Evidence from the lending practices of large and small banks.' *Journal of Financial Economics* 76(2), 237–269

Bloom, Nicholas, Luis Garicano, Raffaella Sadun, and John Van Reenen (2011) 'The Distinct Effects of Information Technology and Communication Technology on Firm Organization.' Available at http://cep.lse.ac.uk/pubs/download/dp0927.pdf

Brynjolfsson, Erik, and Lorin Hitt (2000) 'Beyond Computation: Information Technology, Organizational Transformation and Business Performance.' *The Journal of Economic Perspectives* 14(4), 23–48

Brynjolfsson, Erik, and Shinkyu Yang (1996) 'Information Technology and Productivity: A Review of the Literature.' In *Advances in Computers,* ed. Marvin Zelkowitz, vol. 43 (Elsevier) pp. 179–214

Cadenas, Ximena, and Antoinette Schoar (2011) 'Fighting Procrastination within the Firm.' *NBER Working Paper.* 16944

Cole, Shawn, Martin Kanz, and Leora Klapper (2012) 'Incentivizing Calculated Risk-Taking: Evidence from an Experiment with Commercial Bank Loan Officers'

Dewatripont, Mathias, Ian Jewitt, and Jean Tirole (1999) 'The Economics of Career Concerns, Part I: Comparing Information Structures.' *The Review of Economic Studies* 66(1), 183–198

Einav, Liran, Mark Jenkins, and Jonathan Levin (2012) 'The Impact of Information Technology on Consumer Lending'

Garicano, Luis (2000) 'Hierarchies and the organization of knowledge in production.' *Journal of Political Economy* 108(5), 874–904

Heider, Florian, and Roman Inderst (2012) 'Loan Prospecting.' *Review of Financial Studies* 25(8), 2381–2415

Hertzberg, Andrew, Jose Maria Liberti, and Daniel Paravisini (2010) 'Information and Incentives Inside the Firm: Evidence from Loan Officer Rotation.' *The Journal of Finance* 65(3), 795–828

Holmstrom, Bengt (1979) 'Moral Hazard and Observability.' *The Bell Journal of Economics* 10(1), 74–91

——— (1999) 'Managerial Incentive Problems: A Dynamic Perspective.' *The Review of Economic Studies* 66(1), 169–182

Hubbard, Thomas (2000) 'The Demand for Monitoring Technologies: The Case of Trucking.' *Quarterly Journal of Economics* 115(2), 533–560

Liberti, Jose M., and Atif R. Mian (2009) 'Estimating the effect of hierarchies on information use.' *Review of Financial Studies* 22(10), 4057–4090

Milgrom, Paul, and John Roberts (1990) 'The Economics of Modern Manufacturing: Technology, Strategy, and Organization.' *American Economic Review* 80(3), 511–528

Petersen, Mitchell, and Raghuram Rajan (1995) 'The Effect of Credit Market Competition on Lending Relationships.' *Quarterly Journal of Economics* 110(2), 407–443

Prat, Andrea (2005) 'The Wrong Kind of Transparency.' *The American Economic Review* 95(3), 862–877

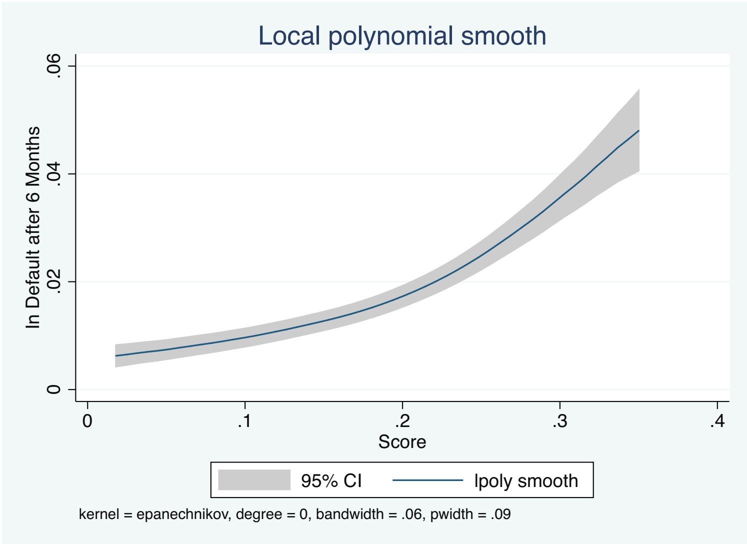
Prendergast, Canice (1993) 'A Theory of Yes Men.' *The American Economic Review* 83(4), 757–770

Rajan, Raghuram (1992) 'Insiders and Outsiders: The Choice between Informed and Arm's-Length Debt.' *The Journal of Finance* 47(4), 1367–1400

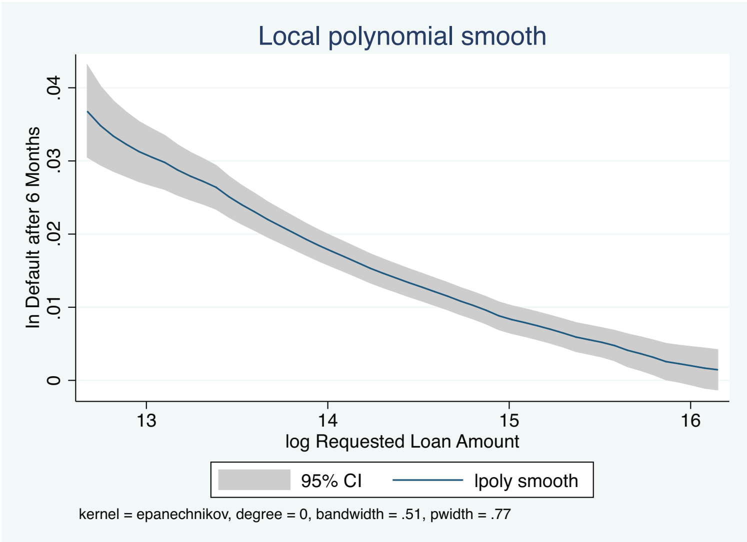
Stein, Jeremy (2002) 'Information production and capital allocation: Decentralized versus hierarchical firms.' *The Journal of Finance* 57(5), 1891–1921

Figure 1: Population Relationships between Default Probability and Credit Scores/Requested Loan Amount

(a) Default Probability, by Score



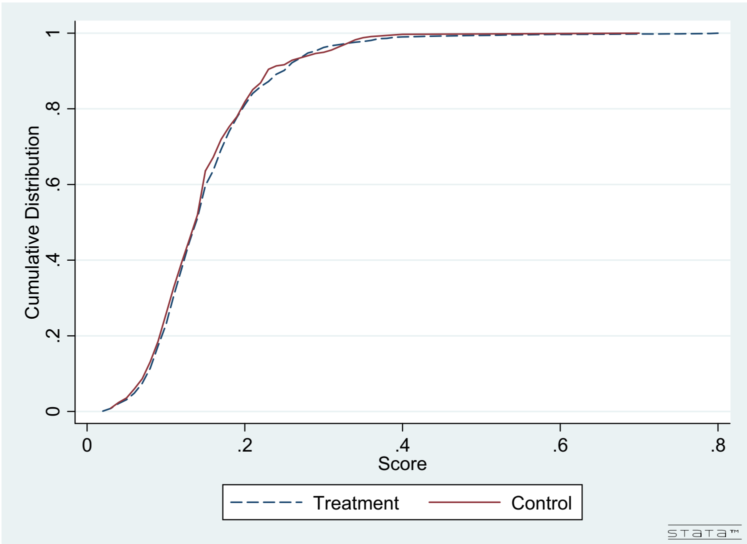
(b) Default Probability, by Requested Loan Amount (log)



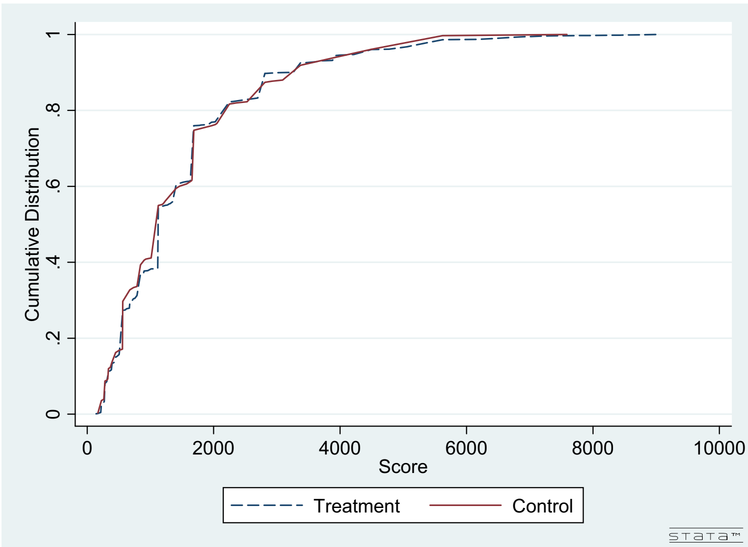
Non-parametric relationship between default probability and (a) credit score, (b) requested loan amount, estimated on the sample of *all* loans approved by BancaMia during October 2010, one month before the roll out of the randomized pilot program.

Figure 2: Cumulative Distributions, by Treatment Group

(a) Scores



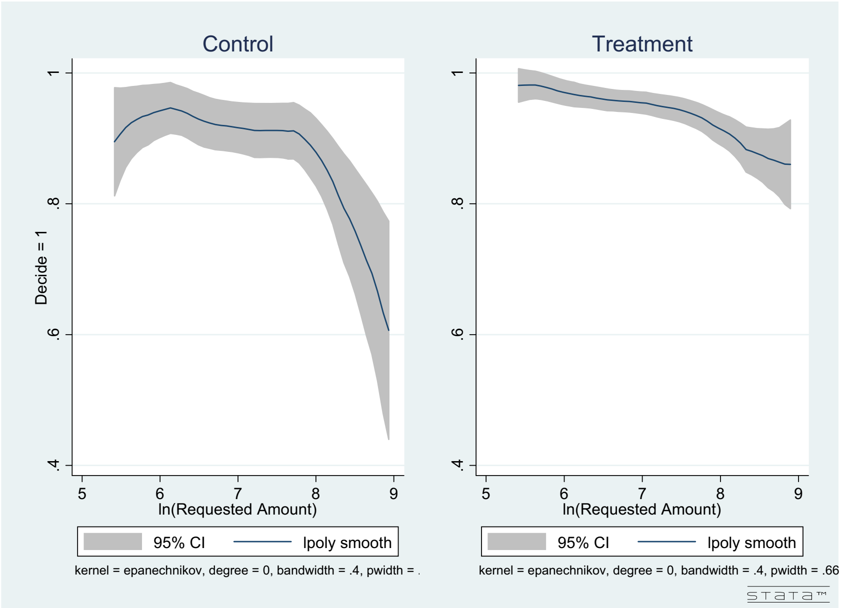
(b) Requested Loan Amounts



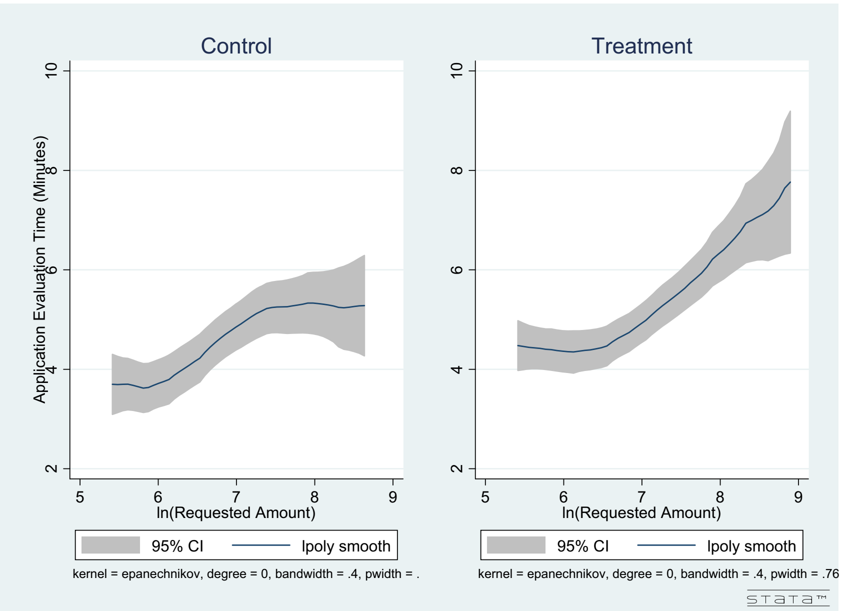
Cumulative Distribution of the scores and requested loan amounts of the loan applications in the randomized pilot program. In the treatment applications, the credit review committee received the score before making final decisions. Scores and requested amounts are pre-determined at the time of the randomization.

Figure 3: Probability of Decision and Evaluation Time, by Application Amount

(a) Probability that Committee Makes Decision



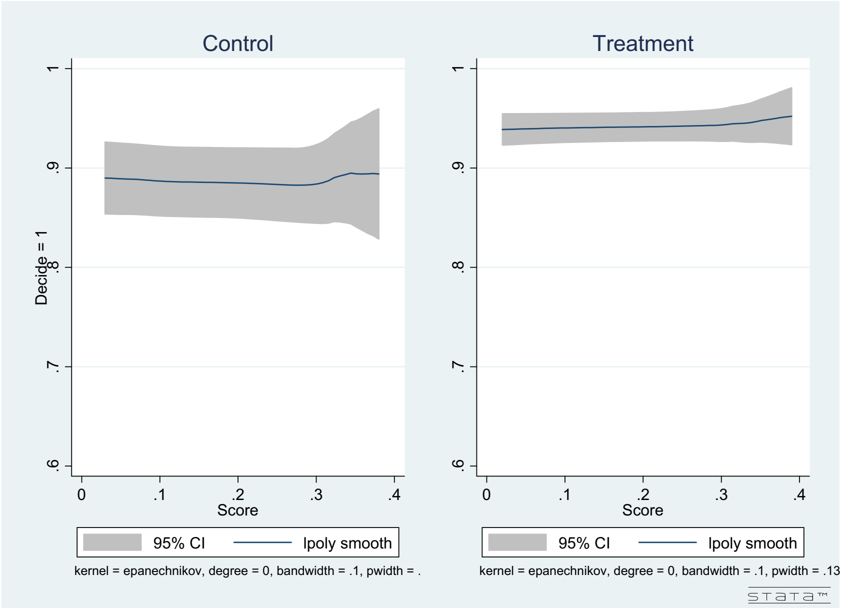
(b) Application Evaluation Time



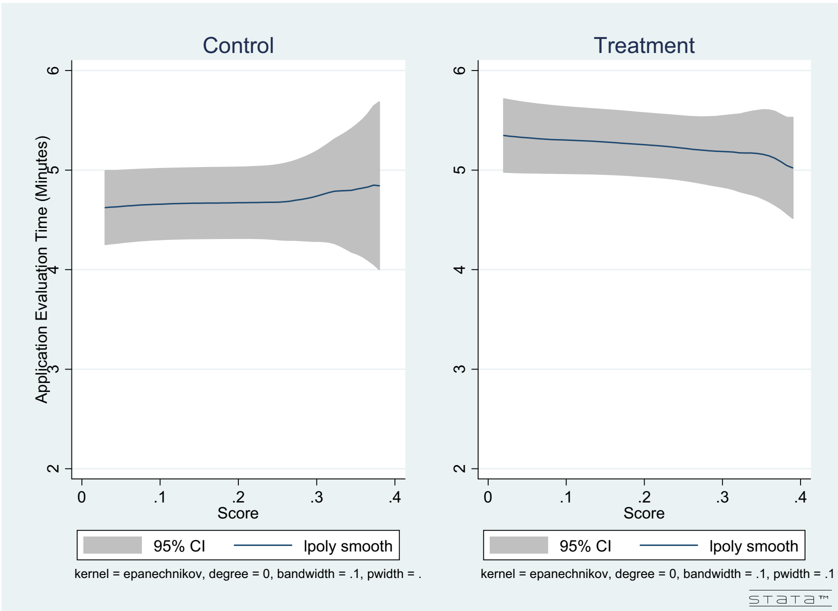
Non-parametric relationship of (a) probability that committee makes a decision on an application (approve or reject) and (b) evaluation time, with application amount.

Figure 4: Probability of Decision and Evaluation Time, by Score

(a) Probability that Committee Makes Decision

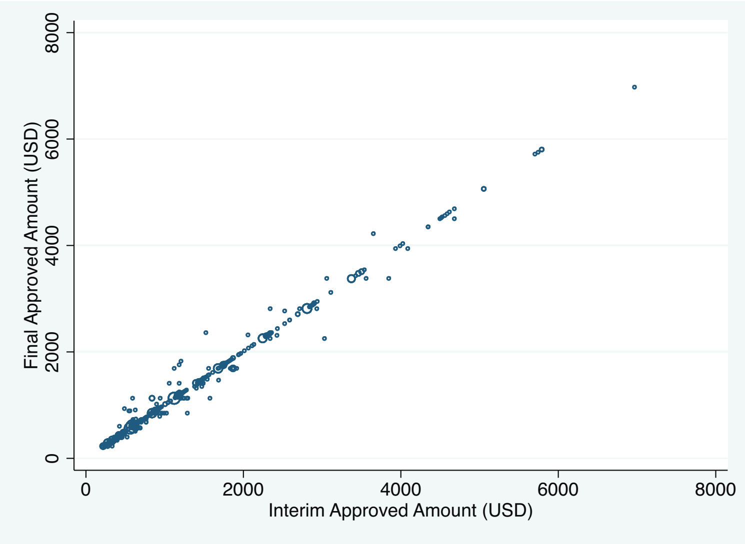


(b) Application Evaluation Time



Non-parametric relationship of (a) probability that committee makes a decision on an application (approve or reject) and (b) evaluation time, with application score.

Figure 5: Approved Loan Amounts before and after Observing Scores (Treatment $T2$)



Plots interim and final approved loan amounts in the subsample of applications in treatment $T2$, in which officers made an interim decision before observing the score and then revised the decision after observing the score.

Table 1: Descriptive Statistics by Randomized Subsample

	(1) Control (n = 335)		(2) Treatment T1 (n = 563)		(3) Treatment T2 (n = 523)		(4) p-value		
	Mean	SD	Mean	SD	Mean	SD	(1) = (2)	(1) = (3)	(2) = (3)
Panel A. Ex Ante Loan Characteristics									
Requested Amount (USD)	1,551.5	1,321.4	1,571.7	1,405.6	1,532.3	1,256.8	0.832	0.852	0.649
Credit Risk Score	0.151	0.068	0.155	0.074	0.158	0.080	0.442	0.184	0.497
First Application (Dummy)	0.146		0.147		0.159		0.962	0.631	0.616
Panel B. Committee Outcomes									
Evalutation Time (minutes)	4.68	3.28	5.13	5.24	5.43	5.34	0.156	**0.021**	0.353
Committee Approves/Rejects (Dummy)	0.890		0.931		0.950		**0.032**	**0.001**	0.221
Panel C. Committee Outcomes, Conditional on Reaching decision									
Loan Approved (Dummy)	0.997		0.987	0.11	0.984	0.13	0.161	0.100	0.717
Panel D. Committee Outcomes, Conditional on Approval									
Amount Approved Requested (Dummy)	0.924		0.945		0.934		0.294	0.663	0.492
Approved Amount/Requested Amount	0.979	0.435	0.975	0.318	0.950	0.293	0.905	0.271	0.187
Abs(Approved Approved - Requested Amount)	266.4	478.8	249.8	484.3	245.6	486.0	0.635	0.557	0.892
Loan Issued (Dummy)	0.754		0.776		0.769		0.487	0.639	0.800
Panel E. Final Outcomes, Conditional on Loan Issued									
Disbursed Amount/Requested Amount	0.959	0.382	0.965	0.297	0.974	0.549	0.828	0.702	0.755
In Default after 6 Months (Dummy)	0.033		0.043		0.036		0.519	0.829	0.618
Defaulted Amount	27.26	166.22	26.43	147.97	35.04	193.26	0.947	0.604	0.476

The last column presents the p-value of a t test of equality of means between the treatment and control applications. The requested amounts in dollars are calculated at prevailing exchange rate of 1,779 pesos per dollar. The credit risk score is a number between zero and one assigned by the credit risk model estimated using BancaMia's historical data on borrower characteristics and repayment performance. The time to decision was calculated from begin and end time of each application's discussion, recorded by the study's research assistants in the field.

Table 2: Descriptive Statistics by Committee Action, Control Group Applications

	Decide (n = 298) (1)		Send Up (n = 16) (2)		More Info (n = 21) (3)	
	mean	sd	mean	sd	mean	sd
Requested Amount (US$)	1,443	1,170	2,480	2,126	2,476	1,994
Credit Risk Score	0.152	0.069	0.155	0.060	0.137	0.047
First Loan (Dummy)	0.154		0.125		0.048	
Time to decision (minutes)	4.608	3.188	5.438	3.405	5.105	4.508
Loan Issued (Dummy)	0.752	0.433	0.750	0.447	0.333	0.483
Disbursed Amount/Requested Amount	0.945	0.272	0.950	0.227	1.486	1.807
In Default after 6 Months (Dummy)	0.031	0.174	0.000	0.000	0.143	0.378

Comparison of application characteristics where the officer reaches a decision—approves or rejects application— (column 1), those where the officer sends the application up for review by the Regional Manager (column 2), and those where the committee decides to send the loan officer out to collect additional information (column 3).

Table 3: Average Treatment Effect of Scores on Committee Output – OLS and LAD

Estimation:	OLS		LAD (Quantile Regression)		
Dependent Variable:	Evaluation Time	Committee Decides	Evaluation time		
			25th %ile	50th %ile	75th %ile
	(1)	(2)	(3)	(4)	(5)
Treatment (T1 and T2)	0.6221**	0.0328*	0.0638	0.3129*	0.5748**
	(0.274)	(0.019)	(0.196)	(0.190)	(0.244)
Treatment (T2)	0.3006	0.0189	0.1600	0.2420	0.1395
	(0.319)	(0.014)	(0.193)	(0.185)	(0.251)
ln(Requested Amount)	1.0114***	-0.0456***	0.4267***	0.5838***	0.7407***
	(0.169)	(0.009)	(0.095)	(0.076)	(0.143)
Credit Risk Score	-1.0271	-0.1381	-0.6988	-1.5684	-1.7864
	(1.429)	(0.112)	(0.885)	(0.978)	(1.543)
First Application	0.7045*	0.0061	0.5239**	0.4478**	0.7891
	(0.390)	(0.018)	(0.204)	(0.175)	(0.504)
Trend	Yes	Yes	Yes	Yes	Yes
Observations	1,405	1,414	1,405	1,405	1,405
R-squared	0.045	0.040			

OLS estimates of the effect of treatment on committee and loan outcomes: Evaluation time in minutes (Column 1), probability that committee reaches decision (column 2), with robust standard errors in parenthesis. LAD estimates of the effect of treatment on evaluation time, with bootstrapped standard errors (500 repetitions) estimated via simultaneous quantile regressions in parenthesis (columns 3 through 5). ***, **, and * indicate significance at the 1%, 5% and 10% levels.

Table 4: Local Average Treatment Effect of Scores on Conditional Outcomes – OLS

Sample Conditioning:	Committee Decides	Committee Approves				Loan Issued		
Dependent Variable:	Committee Approves	ln(Approved Amount)	Approved $\neq$ Requested	| Approved - Requested |	Loan Issued	ln(Issued Amount)	Defaults	Amount Defaulted
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment (T1 and T2)	-0.0112*	0.0110	0.0230	-22.0316	0.0053	0.0190	0.0076	-5.8736
	(0.006)	(0.022)	(0.021)	(31.849)	(0.031)	(0.026)	(0.015)	(13.204)
Treatment (T2)	-0.0026	-0.0226	-0.0121	-11.0843	-0.0064	-0.0115	-0.0087	7.2143
	(0.007)	(0.019)	(0.017)	(28.147)	(0.026)	(0.027)	(0.014)	(12.047)
ln(Requested Amount)	0.0026	0.8753***	0.0099	237.4716***	-0.0362**	0.8231***	-0.0054	5.0767
	(0.004)	(0.010)	(0.011)	(23.269)	(0.015)	(0.016)	(0.007)	(4.403)
Credit Risk Score	-0.1121	-0.5713***	0.2211***	474.6171***	0.0531	-0.5940***	0.4092***	308.9174***
	(0.084)	(0.127)	(0.074)	(164.508)	(0.152)	(0.169)	(0.104)	(81.987)
First Application	-0.0030	0.0002	0.0202	53.0557	0.0255	0.0234	0.0090	8.2727
	(0.009)	(0.024)	(0.019)	(41.561)	(0.032)	(0.028)	(0.018)	(14.700)
Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,319	1,315	1,315	1,315	1,315	1,001	1,001	1,001
R-squared	0.010	0.840	0.009	0.163	0.019	0.777	0.032	0.024

OLS regressions of conditional outcomes on treatment status. Column (1) estimated on subsample of applications where the committee reaches a decision (approves or rejects), columns (2) through (5) estimated on subsample of applications where the committee approved an application, columns (6) through (8) estimated on subsample of issued loans applications. Robust standard errors in parenthesis. ***, **, and * indicate significance at the 1%, 5% and 10% levels.

Table 5: Information Collection and Problem Solving

Committee Choice	Approves (Omitted) (1)	Rejects (2)	More Information (3)	Send to Manager (4)
Score Dummy		1.5592	-0.4521	-0.8127**
		(1.040)	(0.298)	(0.341)
ln(Requested Amount)		-0.2172	0.7647***	0.8350***
		(0.320)	(0.185)	(0.220)
Credit Risk Score		4.9128***	0.8770	3.6230**
		(1.863)	(1.822)	(1.572)
First Application		0.1551	-0.4885	0.2448
		(0.661)	(0.448)	(0.416)
Trend		Yes	Yes	Yes
Observations	1,413			
Pseudo R-squared	0.0727			
Fraction in Control Subsample	0.8866	0.0030	0.0627	0.0478
Marginal Effects:				
Treatment	0.0190	0.0177	-0.0157	-0.0210**
	(0.0180)	(0.0122)	(0.0108)	(0.0092)

Multinomial Logistic Regression estimates of the effect of treatment on final committee actions: make a decision on an application (approve or reject), postpone until the loan officer collects additional information, or send the application to the manager. The first action, make a decision, is the omitted category. The bottom rows present the proportion of each action in the control group and the estimated marginal effect of treatment on the probability that the committee takes an action. Robust standard errors in parenthesis. ***, **, and * indicate significance at the 1%, 5% and 10% levels.

Table 6: Effect of Scores on Overall Output – OLS

Sample Conditioning:	None	Loan Issued				
Dependent Variable: Percentile	Loan Issued (1)	ln(Issued Amount) (2)	Issued $\neq$ Requested (3)	$\vert$ Issued - Requested $\vert$ (4)	Defaults (5)	Amount Defaulted (6)
Treatment (T1 and T2)	-0.0003	0.0216	0.0193	-45.7897	0.0038	-5.8736
	(0.031)	(0.028)	(0.019)	(41.468)	(0.015)	(13.204)
Treatment (T2)	-0.0024	-0.0161	-0.0150	34.2166	-0.0079	7.2143
	(0.026)	(0.026)	(0.017)	(36.529)	(0.013)	(12.047)
ln(Requested Amount)	-0.0441***	0.8194***	0.0115	312.8091***	-0.0060	5.0767
	(0.015)	(0.016)	(0.010)	(31.467)	(0.006)	(4.403)
Credit Risk Score	-0.1080	-0.5831***	0.2172***	572.4381***	0.3851***	308.9174***
	(0.158)	(0.164)	(0.070)	(207.055)	(0.099)	(81.987)
First Application	0.0367	0.0264	0.0209	11.9023	0.0142	8.2727
	(0.032)	(0.027)	(0.018)	(50.113)	(0.018)	(14.700)
Trend	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,414	1,046	1,046	1,046	1,046	1,046
R-squared	0.022	0.771	0.009	0.195	0.030	0.024

OLS estimates of the effect of treatment on overall application outcomes, without conditioning on whether the committee made the decision during the experiment, or the decision was made outside the experiment by either the committee in a later evaluation or by the Regional Manager. Column (1) is estimated on all applications, and columns (2) through (6) on the subsample of applications where the loan was approved. Robust standard errors in parenthesis. ***, **, and * indicate significance at the 1%, 5% and 10% levels.

Table 7: Aggregate Effects on Branch Outcomes

Unit of Observation:	Loan			Branch-Week					
	Score	ln(Requested Amount)	Defaults	ln(Number of Loans)	Score Mean	ln(Sum Requested Amount)	ln(Sum Issued Amount)	Fraction of Loans that Defaults	Fraction of Amount that Defaults
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Experiment Week	-0.0014	0.0026	-0.0014	0.0236	0.0009	0.0481	0.0460	-0.0034	-0.0005
	(0.002)	(0.020)	(0.004)	(0.044)	(0.002)	(0.039)	(0.039)	(0.006)	(0.004)
Branch Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Branch Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	18,327	18,327	18,296	540	540	540	540	540	540
R-squared	0.026	0.019	0.010	0.677	0.361	0.657	0.661	0.261	0.218

OLS regression of loan characteristics on a dummy equal to one if the application was evaluated during a week in which the randomized pilot study was taking place in the branch. Sample contains only approved loans from the eight pilot branches and eight propensity score matched branches and eight propensity-score matched branches (branch matching based on number and total amount of loans approved, average approved loan size and borrower score during October 2010). The sample period is from week 41 of 2010 to week 26 of 2011 (four weeks before and after the pilot program began and ended). Columns 1 through 3 are estimated at the loan level, and 3 through 9 at the branch-week level. Robust standard errors clustered at the branch level in parenthesis. ***, **, and * indicate significance at the 1%, 5% and 10% levels.

Table 8: Information versus Incentives: Effect on Interim and Final Actions in $T2$ – OLS

	Committee Decides		Committee Approves		ln(Approved Amount)		| Approved - Requested |	
Choice:	Interim	Final	Interim	Final	Interim	Final	Interim	Final
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment T2	0.0388**	0.0519***	-0.0151**	-0.0129**	-0.0105	-0.0094	-27.17	-37.76
	(0.019)	(0.018)	(0.007)	(0.006)	(0.023)	(0.023)	(31.34)	(31.85)
ln(Requested Amount)	-0.0488***	-0.0521***	0.0033	0.0034	0.8572***	0.8571***	252.5***	243.8***
	(0.013)	(0.013)	(0.005)	(0.005)	(0.014)	(0.014)	(30.049)	(30.224)
Credit Risk Score	-0.1070	-0.1727	-0.2181*	-0.1939	-0.8223***	-0.8097***	712.8***	693.6***
	(0.135)	(0.134)	(0.130)	(0.129)	(0.177)	(0.177)	-211.2	-208.8
First Application	0.0400*	0.0322	-0.0019	-0.0037	-0.0123	-0.0132	60.19	68.62
	(0.021)	(0.021)	(0.010)	(0.010)	(0.033)	(0.033)	(52.59)	(53.47)
Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	854	854	789	789	784	794	784	794
R-squared	0.042	0.051	0.028	0.024	0.819	0.819	0.183	0.171

OLS estimates the treatment effect on interim committee outcomes before observing the score (odd columns) and on final outcomes after observing the score (even columns). Robust standard errors in parenthesis. ***, **, and * indicate significance at the 1%, 5% and 10% levels.

Table 9: Information versus Incentives: Effect on Interim and Final Actions in $T2$ – ML

Estimation	Interim Outcomes				Final Outcomes			
Action:	Approve (Omitted) (1)	Reject (2)	More Information (3)	Send to Manager (4)	Approve (Omitted) (5)	Reject (6)	More Information (7)	Send to Manager (8)
Treatment T2		1.6285	-0.4361	-0.7142*		1.5914	-0.4621	-1.4662***
		(1.053)	(0.338)	(0.400)		(1.045)	(0.338)	(0.486)
ln(Requested Amount)		-0.3480	0.8792***	0.6257*		-0.3439	0.8894***	0.9451**
		(0.466)	(0.273)	(0.323)		(0.467)	(0.275)	(0.395)
Credit Risk Score		6.7507***	1.3412	2.5523		6.8551***	1.5164	4.8291
		(1.783)	(1.669)	(3.167)		(1.800)	(1.708)	(3.098)
First Application		0.1596	-0.9992	-0.5870		0.1771	-0.9823	-0.4317
		(0.674)	(0.636)	(0.657)		(0.676)	(0.634)	(0.766)
Trend		Yes	Yes	Yes		Yes	Yes	Yes
Observations	853				853			
R-squared								
Pseudo R-squared	0.0818				0.108			
Fraction in Control Subsample	0.8866	0.0030	0.0627	0.0478	0.8866	0.0030	0.0627	0.0478
Marginal Effects:								
Treatment T2	0.0214	0.0165	-0.0160	-0.0219*	0.0343	0.0162	-0.0157	-.0348***
	(0.0207)	(0.0114)	(0.0131)	(0.0129)	(0.0209)	(0.0112)	(0.0129)	(0.0136)

Multinomial Logistic Regression estimates of the effect of treatment on interim committee choices before observing the score (columns 1 through 4) and on final choices after observing the score (columns 5 through 8). The bottom rows present the proportion of each action in the control group and the estimated marginal effect of treatment on the probability that the committee takes an action. Robust standard errors in parenthesis. ***, **, and * indicate significance at the 1%, 5% and 10% levels.

Table A.1: Study Sample: Number of Applications per branch and per Treatment Status

	Control	T1	T2	Total
Branch #:				
1	44	67	62	173
2	89	153	132	374
3	26	51	66	143
4	69	88	87	244
5	18	28	27	73
6	22	26	14	62
7	20	45	38	103
8	47	105	98	250
Total	335	563	524	1,422

Control: the committee makes decision without observing the score. $T1$: the borrower's score is made available at the beginning of the application evaluation. $T2$: the committee makes an interim decision before the score is made available, and the allowed to revise the decision after observing the score.

Table A.2: Effect of Scores on Committee Output, No Controls

Sample Conditioning:	None		Committee Decides	Committee Approved		Loan Issued	
Dependent Variable:	Evaluation Time	Committee Decides	Committee Approves	ln(Approved Amount)	Loan Issued	ln(Issued Amount)	Defaults
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Score Dummy	0.5962**	0.0506***	-0.0113**	0.0281	0.0182	0.0541	0.0099
	(0.242)	(0.019)	(0.005)	(0.050)	(0.028)	(0.056)	(0.014)
Observations	1,412	1,421	1,319	1,315	1,303	1,001	1,001
R-squared	0.003	0.007	0.002	0.000	0.000	0.001	0.000

OLS estimates of the effect of treatment on committee and loan outcomes. Columns (1) and (2) are estimated on all applications, columns (3) and (4) on the subsample of applications where the committee reached a decision, column (5) on the subsample of approved applications, and columns (6) and (7) on the subsample of issued loans. Robust standard errors in parenthesis. ***, **, and * indicate significance at the 1%, 5% and 10% levels.

Table A.3: Interim and Final Decisions in Treatment $T2$

	Final Decision (after Observing Score):				
	Accept Loan	Reject Loan	Obtain More Information	Send Decision to Manager	Total
Interim Decision:					
Accept Loan	482	0	0	1	483
Reject Loan	0	8	0	0	8
Obtain More Information	0	0	20	0	20
Send Decision to Boss	7	0	0	5	12
Total	489	8	20	6	523

Each observation in the matrix represents the two sequential decision made by a committee regarding the *same* application in treatment $T2$. Interim decisions (rows) are the decisions made before observing the score and final decisions (columns) are the revised decisions after observing the score.