USC FBE FINANCE SEMINAR
presented by Antoinette Schoar
FRIDAY, Nov. 15, 2013
10:30 am – 12:00 pm, **Room: JKP-102**

The Incentive Effect of IT: Randomized Evidence from Credit Committees*

Daniel Paravisini — LSE
Antoinette Schoar — MIT

November 6, 2013

Abstract

We distinguish the impact of information technology adoption on information processing costs and agency costs by conducting a randomized control trial with a bank that adopts a new credit-scoring tool. The availability of scores significantly increases credit committees' effort and output on difficult-to-evaluate loan applications. Output increases almost as much in a treatment where the committee receives no new information, but anticipates the score becoming available after it evaluates an application, which suggests that scores reduce incentive problems inside the credit committee. We also show that scores improve efficiency by decentralizing decision-making and equalizing marginal returns across loans.

*We are grateful to Heski Bar-Isaac, Greg Fischer, Luis Garicano, Ulrike Malmendier, Andrea Prat, Anjan Thakor, John Van Reenen and the Arizona State University, Boston University, Columbia University, Darmouth College - Tuck School of Business, HEC School of Management, Imperial College Business School, European Bank for Reconstruction and Development, Federal Reserve Bank of New York, Helsinki Center of Economic Research, Nova School of Business and Economics, Queen Mary School of Economics and Finance, Swiss Finance Institute - USI, University of Illinois Urbana-Champaign, University of Lausanne, University of London - Royal Holloway College, University of Namur, University of Oxford - SAID Business School, University of Toulouse, and Washington University seminar participants; and the CEPR Workshop on Incentives, Management and Organization, CEPR/AMID Development Economics Symposium, the NBER Corporate Finance Workshop, the NBER Productivity, Innovation, and Entrepreneurship Workshop, and the NBER Organizational Economics Workshop participants for helpful comments and discussions. We wish to thank BancaMia for their support, and Gustavo Caballero, Ximena Cadenas, Isabela Echeverry, Santiago Reyes and Lucia Sanchez 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

Technological change has been shown to be associated with broad changes in worker productivity, skill composition, and the wage structure of industries. In particular, the diffusion of information technologies (IT) is accompanied by increases in productivity and the skill premium across firms and industries (Katz and Murphy (1992), Autor, Levy and Murnane (2003), Autor and Dorn (2013)).[1] However, it has been difficult to identify empirically the channel through which IT adoption affects productivity. The identification problem arises from the fact that IT adoption is usually bundled with other organizational innovations, such as changes in job descriptions or compensation structures, which may also affect productivity (Milgrom and Roberts (1990)).[2] The main difficulty in isolating the mechanism lies in the dual role played by most innovations. For example, an IT that allows managers to observe worker output may increase productivity directly by reducing the communication costs between the manager and the worker, or by improving the quality of the manager's decisions. It can also affect productivity indirectly by reducing the cost of monitoring and providing incentives to the worker.

Distinguishing between the information and agency channels is important for understanding the implications of IT adoption on firms' internal organization and boundaries.[3] Existing work attempting to disentangle the two mechanisms using administrative data has had to rely on ad hoc assumptions on whether the information channel or the agency channel will be dominant for any given technology.[4] However, the ex ante classification

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

[2]In addition, firms may adopt ITs in response to changes in the environment that affect productivity directly (see Heider and Inderst (2012) for theoretical application to loan prospecting).

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

[4]For example, Hubbard (2000) identifies two classes of on-board computers in the trucking industry, classifying 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. Bloom, Garicano, Sadun and Van Reenen (2011) classify technologies into communication-enhancing and information-enhancing, although not for the purpose of separating the

relies heavily on the researchers' judgment.

The present paper uses a randomized control trial to empirically identify the causal effect of a particular IT implementation: The adoption of a credit scoring model by a for-profit bank in Colombia specialized in loans to small enterprises. Prior to the adoption of the scoring model, credit committees evaluated loan applications based on raw information about the borrower collected first-hand by loan officers in the field. Committees use the information to decide whether to approve the loan, reject it, or refer it to the regional manager for review. Conditional on approval, committees decide the terms of the loan. The intended goal of the scoring model is to reduce the transaction costs of this decision process. We work with the bank to randomize the roll-out of computer-generated credit scores as an additional input to committees' decision-making process. A credit score is a summary statistic of the borrower's expected default probability, estimated using historical data on loan performance and borrower information in the application.

The credit score may lower the committee's decision-making costs in two ways. On the one hand, it can have a direct effect on the cost of analyzing the information in an application. For example, the scoring model based on population data may produce more precise estimates of the default probability than committee members' judgement, which is based on a small sample of loans. We refer to this mechanism as the *information channel.* On the other hand, scores may affect decision-making indirectly by reducing the cost of incentivizing committee members. For example, by standardizing and making salient the expected quality of an applicant, the score may increase committee members' reputation costs of referring "easy" applications to the boss. We refer to explanations that do not involve the committee obtaining new information from the score as the *agency channel.*

To separate these two channels we set up two treatments. In the first ($T1$), we provide the applicant's score to the credit committee as it sits down to deliberate. In a randomly

information and agency channels.

selected sample of treatment applications ($T2$), 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. Interim decisions in $T2$ allow us to measure the effect of the score *holding constant* the information set under which the committee makes decisions—the agency channel. To guarantee that the information content of the applications is uncorrelated with treatment, we design the trial such that the randomization occurs in real time at the credit committee meeting. This insures that loan officers do not know at the time of collecting the information whether the application will be in the treatment or the control group.[5]

We present three main findings. First, committees are more likely to reach a decision on an application (more output) and spend more time evaluating applications (more effort) when given access to the credit score, as in treatment $T1$. The increase in output and effort is concentrated in marginal—difficult to evaluate—applications, and committees refer to higher level managers half as many applications when they they can see a score. That is, scores induce committees to tackle difficult problems previously solved by managers. This implies that scores and committee effort are complement inputs in the evaluation of marginal applications by committees, and they substitute for managerial inputs in loan production. Second, treatment does not affect average loan amount and default rate, and reduces the variation of marginal realized returns across loans. Thus, treatment improves the cross-sectional allocation of credit without affecting compensation. Together with the first finding, this implies that scores positively impact the productivity of lower level workers in the bank.[6] And third, interim output—decisions made before observing the

[5]Randomizing at the application level comes at the cost of potential spillovers from the treatment to the control group if committee members learn how to estimate scores. Learning spillovers bias the estimates towards zero.

[6]Committees adjust loan size and maturity to mitigate the default risk of more opaque clients. Bank headquarters impose loan interest rates centrally according to product categories. In the appendix, we formally model the optimal credit committee effort and loan size when interest rates are fixed.

score in the second treatment group ($T2$)— also increases relative to the control group. This is consistent with the agency mechanism: interim output increases even though the committee has not received any new information about the borrower. Output increases even further after the committee observes the score, but 75% of the total output increase in $T2$ occurs before the committee observes the score. This implies that the agency channel explains most of the effect of scores on committee output.[7]

Our results present causal evidence on the role of information technologies in shaping the optimal organization of production. Our empirical setting is close to the theoretical setting in Garicano (2000), where workers solve routine problems independently, but ask managers for advice on difficult ones. Our findings provide the first direct corroboration of the main prediction of the model: innovations that lower the information-processing costs for workers lead to more decentralized decision-making inside the firm.[8] The results also provide new evidence related to the ongoing debate on the merits of transparency in agency contexts. In theory, making agents' information observable has an ambiguous effect on the productivity of difficult-to-evaluate workers. On the one hand, observing the agent's information reduces the cost of inducing effort in standard moral hazard (Holmstrom (1979)) or in moral hazard in teams (Dewatripont and Tirole (2005); Visser and Swank (2007)) settings. On the other hand, when agents have career concerns (Holmstrom (1999)), or when they have private information about the productivity of their actions (Prat (2005)), innovations that improve transparency may reduce performance.[9] Consistent with the first view, our findings imply that a technology that reduces information

[7]After observing the score in $T2$, committees never revert a decision to reject a loan and in only one instance they send to the manager a loan after having made an interim decision. This implies that committees take the interim decision seriously despite it being non-binding. This is consistent with committee members caring about their reputation.

[8]In theory, decentralization implies in the long-run a lower unit cost of loan production, a larger optimal span of control, and an increase in total output. However, the empirical approach limits us to studying the short term effects of score adoption, holding loan prices and the size of the firm fixed.

[9]See also Dewatripont, Jewitt and Tirole (1999), and Milbourn, Shockley and Thakor (2001). Performance may decline also if committees disregard their private signals so as to "conform" to the score, as in Prendergast (1993).

asymmetries inside the firm leads to significant performance improvements.

This final conclusion is of key importance in our particular application given the widespread use of credit scoring tools in credit allocation decisions.[10] In an influential paper, Stein (2002) argues that the adoption of technologies that *harden* information in bank lending can lead to larger, more centralized, banks, since these technologies substitute for loan officer discretion and inputs. Our results highlight a countervailing effect: summarizing complex, multidimensional information into a single, easy to interpret statistic, may increase the number and difficulty of tasks that can be delegated to lower-level loan officers by making them easier to monitor and incentivize.

The rest of the paper proceeds as follows. In Section 2, we provide a description of the tasks and incentives of the credit committees and 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, and 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 unsecured loans to micro and small enterprises. The business model of BancaMia, similar to those of other for-profit micro and small business lenders, is based on issuing large numbers of small loans. During October 2010, the month prior to the roll-out of the pilot, the bank issued 20,119 loans totalling $US25.9 million through its 143 branches

[10]Proprietary credit scoring models are commonplace in the banking, credit card, and other consumer lending sectors. There is also a worldwide credit scoring industry for consumer and business lending. FICO, Transunion, Dun and Bradstreet, and Experian are some of the largest industry players. In addition, statistical default models have come under scrutiny by policymakers and academics alike due to their failure to predict the upswing in consumer and mortgage default rates during the 2008 crisis (Rajan, Seru and Vig (2013)).

(average size below $US1,300). The bank makes lending decisions based on information about prospective borrowers collected first-hand 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.[11]

Loan officers 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 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. the borrower's credit score from a private credit rating agency) is put together by the system in a single application file.

The loan prospecting process has three unique features. First, in the absence of financial statements, most field-collected information is "soft", in the sense that it is difficult to verify by anyone other than the loan officer (Stein (2002)). For example, the loan officer may gather the sales figures for a small restaurant by counting the number of tables served during lunch-time and multiplying by the average price of a meal. Second, the loan officer makes an active decision to fill out an application and bring it to the committee for consideration. Thus, applications that reach the committee do not represent the universe of potential borrowers, but only those that pass the initial screening by the officer in the field.[12] Third, officers advice prospective borrowers on how to fill out the application and encourage the riskiest borrowers to request small loan amounts to improve the probability of approval. Thus, application loan amounts do not represent the

[11]BancaMia also 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 in a manner consistent with Rajan (1992) and Petersen and Rajan (1995). However, a very low fraction of the study's sample applications are first-time borrowers, so we ignore this learning aspect of lending in the analysis.

[12]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.

borrower's unconditional demand for credit. 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 the loan prospecting stage, we must consider the possibility that the introduction of scores changes the officers' incentives to gather information and select applicants in the field. 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. In the results section, we also explore how the observable characteristics of the applicant pool change during the pilot week and find little evidence of changes in the prospecting effort.

2.1 Committee Assessment

The application is reviewed by a credit committee composed of the loan officer who collects the information, the branch manager (the loan officer's immediate superior), and one or two additional credit specialists. 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 amount and maturity of the requested loan in order to improve profitability. For example, the committee may decide to approve $US500 for a loan application of $US1,000 if the borrower is deemed to be too unlikely to repay the latter amount. Since rejection rates are extremely low, the main task of the committee is to decide the "correct" loan amount and maturity given the riskiness of the borrower. The committee has no discretion to set the interest rate of loans, which is determined by headquarters based on the type of loan (first-time versus repeat borrower, urban or rural loan). When a committee approves or rejects a loan, we consider that the committee has reached a decision regarding an application.

The third action available to the committee is to refer the application to a regional manager, who evaluates the application and reaches a decision.[13] 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 referrals and additional rounds of information collection represent a substantial cost to the bank in terms of managers' and officers' opportunity cost of time. It is difficult to quantify these costs precisely. The base fixed wage of a regional manager is four to eight times that of a loan officer, which gives a lower bound on the incremental cost of a referral to upper management. Further, since the regional manager must evaluate the application without the officer that collected the information present, she must incur in additional communication costs to access any information not reflected in the application. Referrals imply additional delay costs. This is in part due to the large volume of referred applications and the time constraints of regional managers who 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.[14] Bonuses are calculated on the basis of loans issued, regardless of whether the decision was made by the committee or referred to the manager.[15]

[13]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 the application amount as a control does not change the estimated effect of scores.

[14]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.

[15]There are two potential reasons for compensating committees for loans issued. First, pay based on decisions made penalizes committees for referrals, which may lead to too many bad decisions at the

2.2 Credit Scores

In 2010, BancaMia hired an external consultant to develop a credit risk model that uses the historical information in the credit applications to predict the repayment performance prospective borrowers. The scoring model uses both quantitative (gender, age, location, number of years in business, frequency of late payments in past three years) and qualitative (overall knowledge of business, general sense of the level of organization, quality of information provided, quality of business location, stability and diversity of income) information in the applications. The risk model would be used to produce a credit risk score to be included in the application file. Management's stated objective is 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 score is a proxy for the expected default probability of the loan. Figure 1, panel (a), plots the out-of-sample relationship between the default probability and score in the population of loans issued during October 2010 (the scoring model is calibrated using data for loans issued in 2009). A loan is considered to be in default if interest or principal payments are more than 60 days overdue at six or twelve months after the loan is issued. There is a strong positive association between credit scores and default probabilities. The tight standard error band implies that scores have a good out-of-sample predictive power for future default, and that the data collected by loan officers is informative about borrowers' repayment prospects.

There is a negative relationship between requested loan amounts and default probability in the population (Figure 1, panel (b)). This relationship is consistent with loan officers screening large applications by risky borrowers, or recommending risky borrowers to request smaller loan amounts. This relationship suggests that loan officers form an estimate of the borrower's default probability before bringing the application to the

committee level. Second, pay based on decisions made would eliminate the officer's incentives to monitor the borrower after the loan has been issued.

committee.

3 Trial Design and Descriptive Statistics

We design a randomized controlled trial (RCT) with two goals. The first is to measure the causal effect of scores on committee effort and output, as well as overall loan output and performance. 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 provide a signal of borrower quality that reduces the cost of committee deliberation. For example, the score may reduce the cost of communicating the information collected by the loan officer to the other committee members. Alternatively, the score may reduce the committee's cost of processing the data in the application by assigning population-based weights to borrower characteristics.

In the pure agency mechanism, the score provides a public signal of borrower quality to committee members that was previously only privately observable 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 have a higher tolerance for risk than the bank, or if it takes a substantial loan officer effort to communicate information about a borrower to the committee.

The fundamental distinction between the information and agency mechanisms is that in the latter, the score does not bring new information to the committee: all the information is already in the committee but cannot be used effectively due to an agency conflict. We exploit this difference to isolate 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 implement a four-month pilot program with an RCT design in eight urban branches.[16] In each pilot branch we randomly select the treatment applications for which the committee will be able to see the score of the applicant. Randomization occurs in real time when the committee begins to discuss an application. The committee members are informed of the group assignment at the beginning of the discussion.

We randomize in real time at the application level to guarantee that treatment is orthogonal to applicant characteristics and the information gathering effort by the loan officer. The main disadvantage of our design is the potential for learning spillovers from treatment to control applications. In the extreme case where committee members learn perfectly the algorithm that maps borrower characteristics into scores, then there will be no difference between treatment and control application outcomes. Thus, our estimates can be interpreted as lower bounds on the treatment effect of scores.

In the control group, the committee evaluates the application without observing the score. In the first treatment group ($T1$), the committee receives the score before evaluating 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 evaluates the application and chooses an *interim* action before receiving the score, receives the score after taking the interim action, and then may revise its choice to take a *final* action. When choosing the interim action in $T2$, committees do not have any additional information relative to control applications. However, committee members know that the score will become available after choosing the interim action. Interim actions in treatment $T2$ allow us to measure how committee behavior changes when the information asymmetry between the committee members is expected to decline,

[16] BancaMia has 24 branches. The pilot branches are representative of the average urban branch of the bank. BancaMia also operates rural branches, with a larger fraction of loans associated with agricultural micro-enterprises.

while holding constant the information set available to make the decision. The difference between interim and final actions allow measuring the pure information effect of the score.

The main advantage of the design of $T2$ is that it allows measuring the effect of reducing information asymmetry regardless of the nature of the underlying agency problem between committee members. The interim action is non-binding and unobservable to anyone outside the committee. Thus, the effect of $T2$ on interim actions is designed to capture agency problems inside the committee, and not agency problems between the committee and the regional manager or the rest of the organization. If the main benefit of the score is its information content, then it will be optimal for the committee to devote little effort to the interim decision and wait until the score is revealed. But if the score serves as an incentive device, then output will increase before the score is revealed.

In a short training workshop before the roll-out of the scores, we provide branch managers and loan officers at the eight pilot bank branches with a general explanation of the credit risk model, the scores, the objective of the study (researching the usefulness of the score as an input to the credit evaluation process), and a detailed description of the three treatment groups and the randomization procedure. We report in Appendix Table A.1 the number of applications in each group per branch in the study sample.

3.2 Descriptive Statistics

We present descriptive statistics of the applications in the control and treatment groups in Table 1, as well as the p-values of difference of means tests between the three groups. Predetermined application characteristics—characteristics determined before the randomization takes place—are shown in Panel A. In the control group, average loan amount size is US$1,551, average maturity is 20.9 months, average risk score is 0.151, and the fraction of first-time borrowers is 14.6%. Randomization implies that any differences in predetermined variables between the treatment and control groups are purely by chance.

Table 1 corroborates that randomization created groups that are comparable in terms of pre-determined characteristics, with the only significant difference in means occurring for application maturity, which is one month shorter in treatment group $T1$ than in the control group. Our main specification will include pre-determined characteristics to account for chance differences between groups.[17]

Table 1, Panels B through E, presents the statistics for committee and loan outcomes. Some outcomes, such as the time the committee needs to take an action, are measured for all applications. Others are measured conditional on a particular action of the committee. For example, the approved loan amount is measured conditional 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).[18] 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 indicating that the mean size of approved loans differs little from the mean application amount. Nevertheless, 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. Similar patterns can be found for loan maturity, although the proportion of cases in which the committee modifies the

[17]The application 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. We do not perform a test for maturity because the maturity distribution is highly non-normal, with many observations concentrated in whole year numbers (12, 24, and 36 months).

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

loan application maturity is lower (26.2% of the applications in the control group). The low rejection rate and frequent rate of loan size and loan maturity adjustments suggest that committees decisions occur mostly in the intensive margin, e.g., on how much to lend as opposed to whether or not to lend.

Not all approved loans are issued: only 83.5% of the loans approved during the pilot program appear as issued in the bank's information system. The bank does not record the reason why the borrower decides not to go through with the application. The default rate among the issued loans—fraction of loans more than 30 days late in repayment measured six (twelve) months after the loan was issued—is 3.3% (9.5%) 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 the control. None of the loan characteristics or outcomes conditional on approval is statistically different in the treatment and control groups, except for the average absolute value of the change in maturity.

Table 2 shows the descriptive statistics for applications in the control group conditional on the action taken by the committee—made decision, referred 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 from those referred to the manager (as measured by the score), but have a smaller credit risk than those that required additional information. Committees spend less time evaluating applications where they reach decisions than when they refer an application or 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 ascertain that a decision cannot be reached.

We can track final outcomes for applications when the committee did not make a decision using BancaMia's information system. This allows us 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.[19] Loan outcomes differ substantially depending on the action taken by the committee. For example, loans approved by the manager default with a probability of 8.3% after 12 months and have an approved amount that is 95% of the application amount, while loans approved by the committee after a second round of information collection default with a probability of 13.3% and have an approved amount that is 148% of the application amount.

The complexity of the task of reaching a 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 refer to the manager or collect additional information on the difficult-to-evaluate ones (Garicano (2000)). The statistics in Table 2 suggest that difficult applications take more time to evaluate and that time spent evaluating an application can be used as a measure of effort. Also, the statistics suggest that applications for larger amounts, and with longer maturities, are more difficult to evaluate, while the risk of an applicant (as measured by the score) is uncorrelated with difficulty. We provide in the Appendix a simple model that has these features and can be used as a framework to evaluate the implications of introducing credit scores. The framework predicts that scores will induce the committee to make more decisions, and that the marginal decisions made by the committee are more difficult. We discuss these and other predictions further at the time of presenting the results.

[19]Note, however, that because 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).

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 Section 5):

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

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 $T2$. In some specifications we also include an indicator equal to one if the application is in treatment $T2$ to evaluate differential effects of the two treatments (we do not find any). The vector X_i contains predetermined application characteristics: applicant's credit score, application loan amount, application loan maturity, a dummy if it is the first loan application of the potential borrower, and the date of the application (in weeks).[20]

We begin by presenting the results for outcomes that are measured unconditionally, such as the action taken by the committee or the application evaluation time. The estimated β using these outcomes measures the Average Treatment Effect (ATE) of having a score as an input to the credit evaluation process.

4.1 Committee Actions

We present in Table 3 the estimated effect of introducing a score on the probability that the committee makes a decision (accepts or rejects an application). The point estimate is 4.6 percentage points, statistically significant at the 5% level (column 1). 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 41.8% relative to the baseline proportion

[20] Results without controls are not significantly different, see Appendix Table A.2.

of 11% in the control group. The difference in the effect between $T1$ and $T2$ is positive but not significant (Table 3, column 2).

The data allows identifying two distinct margins through which scores increase committee output: 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 present in in Table 4 the results of estimating a 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.[21]

Treatment has a negative and statistically significant effect on the probability of sending the application to the manager and on the probability of collecting additional information.[22] To evaluate the economic significance of the effects, we report on the bottom rows of Table 4 the implied marginal effect of treatment on the probability of each choice. Observing a score decreases the probability of referring the application to the manager by 2.3 percentage points, a 48% 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 1.7 percentage points, a 27% decline relative to the baseline.

The results suggest that scores reduce the cost of decision-making by committees. Consistent with the prediction in Garicano (2000), lower decision-making costs by committee

[21] We estimate:

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

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.

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

lead to fewer referrals of difficult problems to managers in upper levels of the hierarchy. Scores also substitute for additional costly information collection by loan officers.

4.2 Committee Effort

Our evaluators recorded in real time the beginning and ending time of each application's evaluation. We use the time spent evaluating an application as a measure of committee effort. The estimated effect of introducing a score on the time it takes to evaluate and application is 0.76 minutes, statistically significant at the 1% level (column 3). This implies that committees spend, on average, 16.2% more time per application when scores are available, measured at the mean evaluation time in the control group of 4.7 minutes. Treatment $T2$ has a larger effect on evaluation time than $T1$, but the difference is not statistically significant (column 4). Some 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.

One would conjecture these results are the result of committees making more decisions 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 5 through 10, shows the result of estimating specification (1) using simultaneous quantile regressions for the 25th, 50th, and 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 (columns 5, 7, and 9).

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. In contrast, the additional effect of $T2$ relative to $T1$ on evaluation time appears to be constant across all quartiles (columns 6, 8, and 10). So the additional time spent per application due to the additional decision required in $T2$ does not seem to be related to problem difficulty.

We can further explore the relationship between effort and evaluation difficulty in the data by looking at the cross section of applications. We expect that committees are less likely to make decisions (more likely to ask for help) when applications are more difficult to evaluate. The descriptive statistics in Table 2 suggest that loan application amount is correlated with the probability of making a decision. We confirm this non-parametrically in Figure 2, panel (a): the probability of making a decision drops non-linearly with application amount in the control group applications. The figure also shows that treatment increases the probability of making a decision for the largest applications.

Figure 2, panel (b), explores the cross-sectional patterns in evaluation time by application amount. The non-parametric estimate of the average evaluation time in the treatment applications is above the average for control applications for every application amount. More importantly, evaluation time increases with application amount in the control group, and treatment increases the average time spent on the largest applications. These patterns suggest that application amount is correlated with the difficulty of evaluating an application.

We repeat in Figure 3 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, unconditionally, the *level* of the forecast of the default probability is not correlated with the difficulty of evaluating an application. A potential interpretation of this pattern is that the committee has an unbiased signal about the creditworthiness of the borrower, and the score changes the precision of the signal. The difficulty of evaluating an application is correlated with this precision.

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).[23]

4.3 Loan Outcomes

The only loan outcome that we can measure unconditionally is whether or not the loan was issued. All other outcomes—e.g. amount issued, default probability—are measured only conditionally on the loan being issued. Moreover, when we focus on committee outcomes, the conditioning criterion is narrower—e.g., the probability that the *committee* approves a loan can be measured only if the committee makes a decision, as opposed to referring the application to the manager or postponing the decision to collect additional information.

For outcomes that are measured conditionally, the interpretation of β in specification (1) is complicated by the fact that it captures a combination of two effects: 1) a direct causal effect of treatment on the outcome, and 2) a selection effect driven by the effect of treatment on the conditioning variable. The results so far suggest that the selection

[23] 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.

component may be important when measuring committee outcomes, given that treatment affects the probability of making a decision. We begin by evaluating the effect of treatment without conditioning on the committee's action, and we turn to outcomes conditional on committee action in the next subsection.

In Table 5, we present the estimates obtained from specification (1) using dependent variables that we can measure ex post from the bank's information system. The effect of scores on the probability that the loan is issued is close to zero and not statistically significant (Table 5, column 1). This implies that the addition of scores to the loan production process does not affect the overall extensive margin of lending. This also means that the selection component of the β estimates for outcomes measured conditional on the loan being issued (loan amount, default) is negligible.

Scores do not have a statistically significant effect on the average level of any of the measured outputs: loan size, probability that loan amount and application amount are different, absolute value of the loan amount adjustment, and default probability (Table 5, columns 2 through 6). Note that if banks have an unbiased signal of the creditworthiness of the borrower and scores increase the precision of the this signal, the effect of the score on loan amount and default are ambiguous. Banks will lend less to some borrowers and more to others, and this will increase the default probability of some borrowers and lower it for others. The net effect will depend on the distribution of borrower characteristics in the population.

This reallocation across applications should have, in contrast, a strong effect on the relationship between loan contract characteristics and default probability in the cross-section. In the simple framework provided in the Appendix, we show that in the extreme case where the bank has a zero-variance signal of the borrower's creditworthiness (measured as the sensitivity of the borrower's default probability to loan size), the cross-sectional relationship between loan size and default probability becomes flat. The reason is that without any uncertainty about the borrower's creditworthiness, the optimal cross-sectional allocation is the one that equalizes the marginal expected return across borrowers, which in turn implies that loan sizes do not predict default probabilities in the cross-section (conditional on the information set of the bank). On the other hand, when there is uncertainty about the borrower's type, this relationship depends on how the precision of the signal varies in the cross-section of borrowers.

There is a strong cross-sectional correlation between loan amount and default in the control group applications. The slope of loan amount is negative, and on loan maturity is positive in a linear probability model after conditioning for pre-determined observable characteristics (see Appendix Table A.3). In Table 6 we document how these slopes change with treatment. Both slopes become flatter: the slope on loan amount increases by 0.073 (from -0.262) and the slope on maturity decreases by 0.119 (from 0.094). Thus, treatment reduces the cross-sectional correlation between loan size and default by 28% and that between loan maturity and default probability essentially disappears. This implies that the addition of scores to the loan production process substantially decreases the equilibrium cross-sectional correlation between loan contract characteristics that potentially affect default, and ex post default probabilities. This is consistent with the interpretation that scores reduce the uncertainty about the borrower's creditworthiness.

4.4 Committee Conditional Output

To illustrate the complication introduced by measuring the effect of treatment on outcomes that are conditional on committee actions, consider the case of loan size approved by the committee. This loan size can be measured conditional on the committee approving an application, as opposed to rejecting it, sending it to the manager, or postponing the decision until more information is collected. Scores may have a direct effect on approved loan size, holding constant the set of applications. This is the Local Average Treatment

Effect (LATE) of scores on loan amount. Scores also change the set of applications that the committee decides on. These marginal applications are likely to be different along dimensions that are correlated to loan size. In fact, we already documented that this is the case: marginal applications are for larger loan amounts and more difficult to evaluate than the infra-marginal ones. Thus, treatment changes the composition of applications approved by the committee in a way that will likely affect average loan size even if the LATE is zero.

Disentangling the Local Average Treatment and selection effects is typically difficult without an additional instrumental variable for the selection effect.[24] Our setting, however, provides a unique advantage: we can evaluate outcomes of marginal committee decisions due to selection, because these decisions are made regardless by either the manager or by the committee in a second evaluation. To follow our example above, suppose we find that committees approve loans that are significantly larger when the score is available. We know from the results in Table 5 that treatment does not affect average final loan size. This implies that an observed change in loan size approved by the committee is due to selection: treatment induces committees to decide on larger applications that would have otherwise been referred to the manager or decided after collecting more information. In fact, since we found that treatment does not have any effect on the average level of final outcomes, *any* difference that we find between treatment and control groups at the committee level is due to selection.

We present the estimates of β using outcomes conditional on committee actions in Table 7, which includes the conditioning variable in the first row. Conditional on making a decision, the probability that a committee rejects an application increases by 0.9 percentage points in the presence of scores, significant at the 10% level (column 1). This estimate implies a three-fold increase in the proportion of applications rejected by the

[24] See Lee (2008) for a recent discussion.

committee relative to the baseline probability of 0.3% in the control group. Moreover, assuming that all the additional rejections come from the marginal decisions, the estimate implies that committees reject 13% of the marginal cases they decide on when scores are used as an input ($(0.9-0.3)/4.6$). Given that the overall probability that the loan is issued is unchanged, the entire increase must come from applications that would have been rejected by the manager or by the committee in a second decision.

Conditional on the committee approving the loan, scores do not have a significant effect on other committee outcomes. So even though committees are deciding on a larger proportion of marginal cases when the score is available, the average credit supply and loan maturity do not change, and neither do the frequency or amount of revisions to the requested amounts and maturities.

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.

In this subsection, we investigate 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. We use the following difference-in-differences 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 8, columns 1 through 4, estimated using all the loans approved between week 41 of 2010 (four weeks before the pilot starting date) and week 26 of 2011 (four weeks after the pilot end date). 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. In column 5 of Table 8, we test whether default rates are better predicted by the score during pilot weeks by augmenting the default regression with the score interacted with pilot week as a right-hand side variable. The interaction coefficient is not significantly different from zero, indicating that the predictive power of the score does not change during the pilot weeks. These results imply that the introduction of scores did not affect the applicant pool or the quality of the information collected by the loan officer.

We can also use the difference-in-differences approach to evaluate whether the pilot affected total branch output —subject to the caveat that we did not randomize neither the selection of the branches not the timing of the experiment. We estimate specification

(3) aggregated at the branch-week level and present the results in Table 8, columns 5 through 9. We report estimates using the (log) number of loans issued, the (log) sum of requested and approved amounts, the fraction of loans that default, and the fraction of debt that defaults (value-weighted defaults) as dependent variables. The point estimates on the number and amount of lending are positive, and those on default (unweighed and value-weighted) are negative, but only the coefficient on the fraction of loans that defaults is 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, the results are lower bound estimates on the long run effect of scores on total output.

4.6 Discussion

The results in this section 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 enable committees to reach decisions on applications that are more difficult to evaluate. Despite the upward shift in the difficulty of the tasks performed, the average loan amount, maturity, and default rates remain unaltered.

The addition of scores has first-order consequences on the reallocation of credit across borrowers. The evidence suggests that capital is reallocated in a way that equalizes the expected marginal return across loans, which is consistent with the scores reducing the uncertainty about borrower quality.

The increased output by committees substitutes for other, more expensive, inputs to the production of loans, in particular the time of higher-level managers who are more expensive for the bank. This implies that scores increase the decentralization of the decision-making process of the bank and reduce the managers' workload.

By design, the pilot trial holds constant interest rates, managers' span of control, committee members' compensation, and other dimensions of the loan production process. However, the new equilibrium level for these variables is likely to change after score adoption (interest rates may drop to reflect the lower production costs, managers may oversee more branches, etc.). Thus, the estimated effects on output might not capture fully the general equilibrium impact of scores. Changes along these dimensions are likely to induce further improvements in productivity.

5 The Information and Agency Channels

The results presented so far measure the effect of scores on the final decisions made by the committee and the manager. In this section, we turn our attention towards 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 interim 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 the action is chosen.

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 into 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 scope for 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.

The results in the previous section allow ruling out some interpretations. In particular, since we do not observe an effect on average lending or defaults, scores cannot be debiasing the assessments made by officers. If scores reduce agency costs, it is not because officers are systematically underestimating or overestimating the default probability. Rather, it is because officers provide signals that are too noisy relative to the first best. This would occur if producing a precise signal requires costly effort, or if the loan officer is more risk averse than the bank, and prefers to overstate the uncertainty about borrower quality to reduce the variance in her compensation. The framework in the Appendix is based on this observation.

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 modifying 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, 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 the results in Table 9. 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 coefficient is 0.039: the probability the a committee makes decision increases by 3.9 percentage points before observing the score in $T2$. The magnitude is smaller than that of the effect 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 score. In addition, the expectation of receiving a score significantly reduces the probability that committees refer an application to the manager in the interim decisions (see Appendix Table A.5).

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, 1.3 percentage points (column 3), is larger than the increase in the final outcomes, 1.1 percentage points (column 4), although again, the estimates are not statistically distinguishable. Appendix Table A.4 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.

Two conclusions can be drawn from these results. First, the bulk of the effect of scores on committee output occurs even while holding committees' information set constant. This is consistent with the agency mechanism: scores induce committees to make more decisions. Second, the pure information effect of scores is small relative to the agency effect. 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

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 committees' overall output and reduce the need for higher-level manager involvement in the decision-making process. Thus, the paper presents direct evidence on how information technologies can lead to the decentralization of decision-making 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, we evaluate how committee decisions change in a treatment that makes the score available to all committee members soon after they have reached a decision. We find that the threat of making scores available after making a decision increases committee output. 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.[25] In our context, the supervisors are able to observe loan officers' choices even in the absence of the score, for example, when they review the loan officers' performance and bonus payments on a quarterly basis. Thus, scores increase the *immediacy* and *ease* which the principal can monitor the agents but not *whether* they get reviewed. Scores also affect 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 induce 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

Autor, David, and David Dorn (2013) 'The Growth of Low-Skill Service Jobs and the Polarization of the U.S. Labor Market.' *The American Economic Review*

Autor, David, Frank Levy, and Richard Murnane (2003) 'The Skill Content of Recent Technological Change: An Empirical Exploration.' *Quarterly Journal of Economics* 118(4), 1279–1334

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

[25]Note, in contrast, that our results have nothing to say about IT based solutions that completely substitute committee judgement. Berg, Puri and Rocholl (2013) show evidence that piece rate incentives and a lending rule solely based on a credit scoring model leads to loan officers to distort the information that is entered into the model.

Berg, Tobias, Manju Puri, and Jorg Rocholl (2013) 'Loan Officer Incentives and the Limits of Hard Information.' Available at http://

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, and Jean Tirole (2005) 'Modes of Communication.' *Journal of Political Economy* 113(6), 1217–1238

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

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

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

Katz, Lawrence, and Kevin Murphy (1992) 'Changes in Relative Wages, 1963-1987: Supply and Demand Factors.' *Quarterly Journal of Economics* 107(1), 35–78

Lee, David (2008) 'Training, wages, and sample selection: Estimating sharp bounds on treatment effects.' *Review of Financial Studies* 76(3), 1071–1102

Milbourn, Todd, Richard Shockley, and Anjan Thakor (2001) 'Managerial career concerns and investments in information.' *RAND Journal of Economics* 32(2), 334–351

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

Rajan, Raghuram, and Julie Wulf (2006) 'The Flattening Firm: Evidence from Panel Data on the Changing Nature of Corporate Hierarchies.' *The Review of Economics and Statistics* 88(4), 759–773

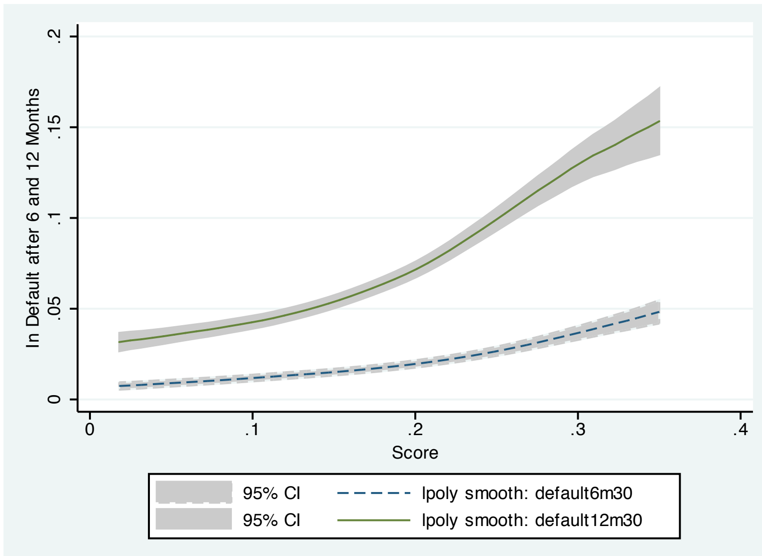
Rajan, Uday, Amit Seru, and Vikrant Vig (2013) 'The Failure of Models that Predict Failure: Distance, Incentives and Defaults.' *Journal of Financial Economics*

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

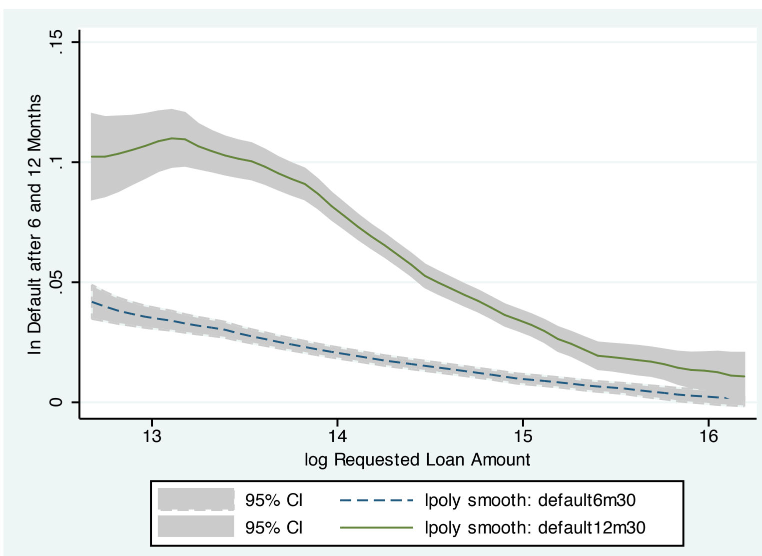
Visser, Bauke, and Otto Swank (2007) 'On Committees of Experts.' *Quarterly Journal of Economics* 221(1), 337–372

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

(a) Default Probability, by Score



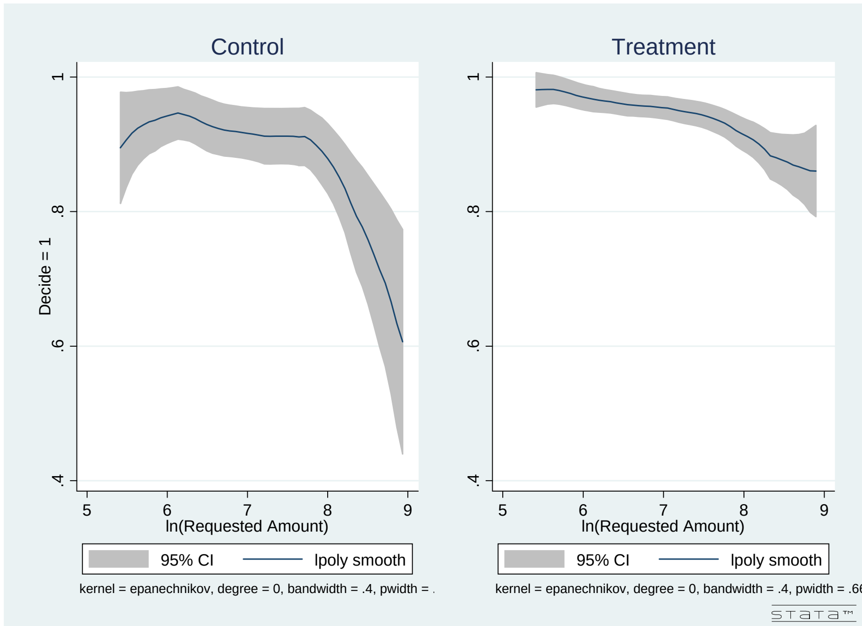
(b) Default Probability, by Requested Loan Amount (log)



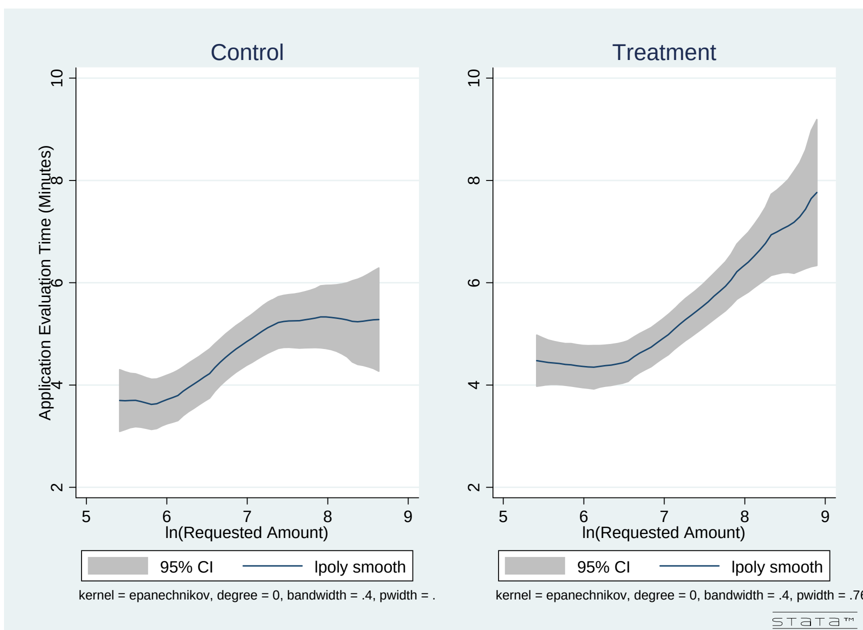
Non-parametric relationship between 6-month and 12-month default probabilities and (a) credit score, (b) requested loan amount, estimated on the sample of *all* loans approved by BancaMia during October 2010, one month before the roll-out of the randomized pilot program.

Figure 2: Probability of Decision and Evaluation Time, by Application Amount

(a) Probability that Committee Makes Decision



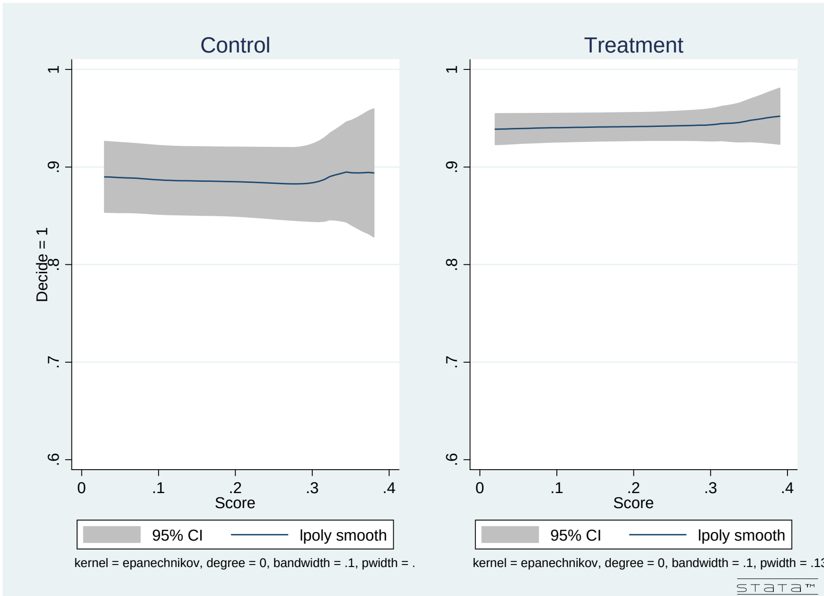
(b) Application Evaluation Time



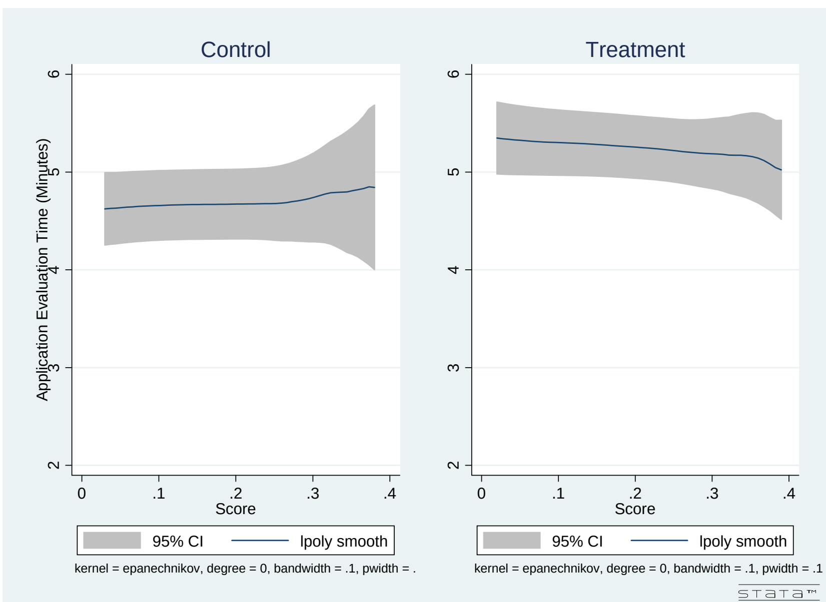
Non-parametric relationship of (a) probability that committee makes a decision on an application (approve or reject) and (b) evaluation time, with application amount.

Figure 3: 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.

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 Application Characteristics									
Application 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
Application Maturity (Months)	20.9	9.8	22.0	10.4	22.1	10.4	0.109	**0.086**	0.880
Credit Risk Score	0.151	0.068	0.155	0.074	0.158	0.080	0.459	0.201	0.512
First Application (Dummy)	0.146	0.354	0.147	0.355	0.159	0.366	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	0.314	0.931	0.254	0.950	0.218	**0.032**	**0.001**	0.221
Panel C. Committee Outcomes, Conditional on Reaching decision									
Loan Approved (Dummy)	0.997	0.058	0.987	0.11	0.984	0.13	0.161	0.100	0.717
Panel D. Committee Outcomes, Conditional on Approval									
Amount Approved $\neq$ Application (Dummy)	0.698	0.460	0.737	0.441	0.692	0.462	0.230	0.849	0.108
Approved Amount/Application Amount	0.979	0.435	0.975	0.318	0.950	0.293	0.905	0.271	0.187
|Approved Approved - Application Amount|	266.4	478.8	249.8	484.3	245.6	486.0	0.635	0.557	0.892
Maturity Approved $\neq$ Application (Dummy)	0.262	0.440	0.278	0.449	0.307	0.462	0.609	0.174	0.314
Approved Maturity/Application Maturity	0.985	0.290	1.000	0.264	0.983	0.371	0.471	0.922	0.404
|Approved Maturity - Application Maturity|	2.3	4.7	2.4	5.0	3.2	6.0	0.616	**0.023**	**0.032**
Loan Issued (Dummy)	0.835	0.372	0.855	0.353	0.840	0.367	0.447	0.840	0.524
Panel E. Final Outcomes, Conditional on Loan Issued									
Disbursed Amount/Application 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)	0.036	0.188	0.039	0.193	0.037	0.190	0.881	0.954	0.917
In Default after 12 Months (Dummy)	0.095	0.293	0.088	0.283	0.089	0.284	0.757	0.781	0.976
Defaulted Amount (6 months)	27.26	166.22	26.43	147.97	35.04	193.26	0.947	0.604	0.476
Defaulted Amount (12 months)	62.67	238.07	71.84	257.73	74.11	265.39	0.650	0.584	0.902

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

	Decide (n = 298) (1)		Send Up (n = 16) (2)		More Information (n = 21) (3)	
	mean	sd	mean	sd	mean	sd
Panel A. Ex Ante Application Characteristics						
Application Amount (US$)	1,443	1,170	2,480	2,126	2,476	1,994
Application Maturity (Months)	20.3	9.3	26.3	12.2	25.1	13.3
Credit Risk Score	0.152	0.069	0.155	0.060	0.138	0.046
First Loan (Dummy)	0.154		0.125		0.048	
Panel B. Outcomes						
Time to decision (minutes)	4.608	3.188	5.438	3.405	5.105	4.508
Loan Issued (Dummy)	0.832		0.750		0.714	
Amount Approved $\neq$ Application (Dummy)	0.924		1.000		1.000	
Approved Amount/Application Amount	0.945	0.272	0.950	0.227	1.486	1.807
|Approved Approved - Application Amount|	287.4	499.1	262.9	309.8	1477.0	2153.0
In Default after 6 Months (Dummy)	0.028		0.000		0.200	
In Default after 12 Months (Dummy)	0.093		0.083		0.133	
Defaulted Amount (after 6 Months)	25.8	164.0	0.0	0.0	121.4	321.2
Defaulted Amount (after 12 Months)	66.4	246.6	28.6	99.2	0.0	0.0

Comparison of application characteristics where the officer reaches a decision—approves or rejects application—(column 1), those where the officer refers the application to the regional manager (column 2), and those where the committee decides to collect additional information (column 3).

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

Estimation:	OLS		OLS		LAD (Quantile Regression)					
Dependent Variable:	Committee Decides		Evaluation Time		Evaluation time					
					25th %ile		50th %ile		75th %ile	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment (T1 and T2)	0.046**	0.038**	0.760***	0.624**	0.183	0.123	0.426***	0.383**	0.663***	0.662***
	(0.018)	(0.019)	(0.229)	(0.273)	(0.192)	(0.194)	(0.157)	(0.193)	(0.231)	(0.254)
Treatment (T2)		0.016		0.283		0.057		0.133		0.113
		(0.013)		(0.321)		(0.170)		(0.171)		(0.249)
ln(Requested Amount)	-0.032***	-0.032***	0.751***	0.747***	0.084	0.053	0.239*	0.238	0.664***	0.617***
	(0.012)	(0.012)	(0.226)	(0.226)	(0.145)	(0.140)	(0.133)	(0.145)	(0.183)	(0.185)
ln(Requested Maturity)	-0.026	-0.026	0.653**	0.656**	0.676***	0.692***	0.680***	0.670***	0.227	0.265
	(0.018)	(0.018)	(0.332)	(0.330)	(0.222)	(0.222)	(0.208)	(0.225)	(0.308)	(0.303)
Credit Risk Score	-0.105	-0.107	-1.28	-1.301	-1.181	-1.175	-1.970**	-1.464	-1.761	-1.971
	(0.111)	(0.111)	(1.457)	(1.458)	(0.752)	(0.818)	(0.945)	(1.077)	(1.564)	(1.558)
First Application	0.009	0.009	0.695*	0.692*	0.408**	0.426**	0.541***	0.571***	0.764	0.757
	(0.018)	(0.018)	(0.387)	(0.388)	(0.187)	(0.194)	(0.186)	(0.204)	(0.509)	(0.535)
Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,406	1,406	1,397	1,397	1,397	1,397	1,397	1,397	1,397	1,397
R-squared	0.041	0.042	0.048	0.049						

OLS estimates of the effect of treatment on committee and loan outcomes: probability that committee reaches decision (column 1), evaluation time in minutes (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: Costly Information and Referrals

Committee Choice	Approves (Omitted) (1)	Rejects (2)	More Information (3)	Send to Manager (4)
Treatment (T1 and T2)		1.3236	-0.5439*	-0.9038***
		(1.049)	(0.305)	(0.344)
ln(Application Amount)		0.0851	0.7971***	0.1790
		(0.466)	(0.243)	(0.308)
ln(Application Maturity)		-0.2176	-0.1358	1.7474***
		(0.731)	(0.386)	(0.570)
Credit Risk Score		4.8843**	0.2146	2.8701*
		(2.121)	(2.001)	(1.537)
First Application		0.4442	-0.6193	0.2574
		(0.673)	(0.487)	(0.419)
Trend		Yes	Yes	Yes
Observations	1405			
Pseudo R-squared	0.0875			
Fraction in Control Subsample	0.8866	0.0030	0.0627	0.0478
Marginal Effects:				
Treatment	0.0281	0.0124	-0.0174*	-0.0231**
	(0.0166)	(0.0100)	(0.0104)	(0.0094)

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 (referrals). 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 5: Effect of Scores on Overall Output – OLS

Sample Conditioning:	None	Loan Issued				
Dependent Variable:	Loan	ln(Issued	Issued $\neq$	|Issued -	In Default after	
Percentile	Issued	Amount)	Application	Application|	6 months	12 months
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment (T1 and T2)	0.0020	0.0107	-30.1268	0.0107	-0.0038	-0.0158
	(0.024)	(0.018)	(38.713)	(0.025)	(0.013)	(0.020)
ln(Application Amount)	-0.0018	-0.0082	307.2***	0.7760***	-0.0036	-0.0327**
	(0.020)	(0.015)	(49.378)	(0.029)	(0.009)	(0.013)
ln(Application Maturity)	-0.0314	0.0445*	14.51	0.0987**	-0.0075	0.0229
	(0.032)	(0.026)	(59.88)	(0.042)	(0.015)	(0.021)
Credit Risk Score	0.0603	0.2064***	568.5***	-0.6069***	0.3256***	0.6472***
	(0.142)	(0.068)	(207.829)	(0.165)	(0.086)	(0.135)
First Application	0.0421	0.0198	12.2255	0.0244	0.0087	-0.0197
	(0.026)	(0.018)	(50.211)	(0.027)	(0.016)	(0.021)
Trend	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,406	1,048	1,048	1,046	1,165	1,165
R-squared	0.007	0.013	0.195	0.773	0.022	0.043

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 6: Effect on Relationship between Default and Contract Characteristics – OLS

Dependent Variable:	In Default after	
	6 months	12 months
	(1)	(2)
Treatment (T1 and T2)	0.0409	-0.1752
	(0.091)	(0.171)
ln(Loan Amount)	-0.0948***	-0.2619***
	(0.031)	(0.051)
ln(Loan Amount) × Treatment	0.0209	0.0733**
	(0.020)	(0.035)
ln(Loan Maturity)	0.0401	0.0944*
	(0.041)	(0.056)
ln(Loan Maturity) × Treatment	-0.0628*	-0.1190**
	(0.033)	(0.050)
ln(Application Amount)	0.2993***	0.5325***
	(0.084)	(0.129)
ln(Application Maturity)	0.0655**	0.1477***
	(0.026)	(0.042)
Credit Risk Score	0.0004	0.0236
	(0.035)	(0.043)
First Application	0.0037	-0.0221
	(0.017)	(0.023)
Trend	Yes	Yes
Observations	1,100	1,100
R-squared	0.047	0.091

OLS estimates of the effect of treatment on the conditional correlation between loan contract characteristics and loan repayment. Robust standard errors in parenthesis. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table 7: Scores and Conditional Committee Outcomes – OLS

Sample Conditioning:	Committee Decides	Committee Approves					
		Loan Amount			Loan Maturity		
Dependent Variable:	Committee Approves Dum.	Approved (log USD)	Approved $\neq$ Application	|Approved - Application|	Approved (log Months)	Approved $\neq$ Application	|Approved - Application|
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment (T1 and T2)	-0.0092*	-0.0001	0.0322	-29.2229	0.0282	0.0197	0.3434
	(0.005)	(0.020)	(0.030)	(28.585)	(0.029)	(0.019)	(0.317)
ln(Application Amount)	-0.0006	0.8752***	0.0357	212.5798***	0.0092	0.1075***	-0.0333
	(0.005)	(0.015)	(0.022)	(28.189)	(0.020)	(0.028)	(0.316)
ln(Application Maturity)	0.0016	-0.0002	-0.0360	59.7938**	0.0382	-0.3075***	2.1669***
	(0.009)	(0.024)	(0.036)	(29.936)	(0.034)	(0.053)	(0.540)
Credit Risk Score	-0.0886	-0.5715***	0.5898***	452.2661***	0.4515**	-0.0730	5.4257**
	(0.082)	(0.128)	(0.168)	(165.744)	(0.180)	(0.099)	(2.131)
First Application	-0.0056	-0.0002	0.0230	51.1913	-0.0228	0.0212	-0.1109
	(0.009)	(0.024)	(0.034)	(41.627)	(0.034)	(0.022)	(0.422)
Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,315	1,314	1,314	1,314	1,314	1,314	1,314
R-squared	0.007	0.840	0.022	0.165	0.010	0.136	0.051

OLS regressions of conditional outcomes on treatment status. Column (1) is estimated on the subsample of applications where the committee reaches a decision (approves or rejects), columns (2) through (7) are estimated on the subsample of applications where the committee approved an application. Robust standard errors in parenthesis. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table 8: Aggregate Effects on Branch Outcomes - Difference-in-Differences Estimate

Unit of Observation:	Loan					Branch-Week			
	Score	ln(Application Amount)	ln(Issued Amount)	In Default after 12 Months	In Default after 12 Months	ln(Number of Loans)	ln(Sum Application Amount)	Fraction of Loans that Defaults	Fraction of Amount that Defaults
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Pilot Week	-0.0014	-0.0194	-0.0185	-0.0065	0.0071	-0.0273	-0.0074	-0.0121*	0.0007
	(0.002)	(0.020)	(0.019)	(0.007)	(0.013)	(0.044)	(0.047)	(0.006)	(0.005)
Risk Score					0.6949***				
					(0.039)				
Pilot Week x Risk Score					-0.0812				
					(0.082)				
Branch Dummies	Yes	Yes		Yes	Yes	Yes	Yes	Yes	Yes
Week Dummies	Yes	Yes		Yes	Yes	Yes	Yes	Yes	Yes
Branch Trends	Yes	Yes		Yes	Yes	Yes	Yes	Yes	Yes
Observations	18,270	18,270	18,270	18,270	18,270	525	525	525	525
R-squared	0.044	0.028	0.033	0.016	0.016	0.754	0.699	0.283	0.170

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 (branch matching based on number and total amount of loans approved, average approved loan size, and borrower score measured in 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 9: Information versus Incentives: Effect on Interim and Final Actions in $T2$ – OLS

Outcome:	Committee Decides Dummy		Committee Approves Dummy		ln(Approved Amount)		Approved $\neq$ Application Dummy		|Approved - Application|	
Choice:	Interim	Final	Interim	Final	Interim	Final	Interim	Final	Interim	Final
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment T2	0.0388**	0.0524***	-0.0136**	-0.0113*	-0.0023	-0.0064	0.0060	0.0016	-28.8	-39.73
	(0.019)	(0.018)	(0.007)	(0.006)	(0.023)	(0.022)	(0.029)	(0.034)	(31.43)	(31.93)
ln(Application Amount)	-0.0257	-0.0378**	0.0013	0.0012	0.746***	0.776***	0.0028	0.0640**	288.5***	285.1***
	(0.016)	(0.015)	(0.007)	(0.007)	(0.022)	(0.021)	(0.026)	(0.029)	(42.20)	(42.50)
ln(Application Maturity)	-0.0572**	-0.0352	-0.0015	-0.0012	0.280***	0.237***	0.0397	-0.0399	-90.5*	-103.4*
	(0.024)	(0.023)	(0.015)	(0.015)	(0.036)	(0.035)	(0.041)	(0.047)	(52.80)	(53.20)
Credit Risk Score	-0.0925	-0.1607	-0.1877	-0.1632	-0.821***	-0.762***	0.456**	0.549**	708.3***	690.4***
	(0.132)	(0.132)	(0.129)	(0.129)	(0.155)	(0.152)	(0.212)	(0.238)	(204.20)	(201.30)
First Application	0.0402*	0.0318	-0.0034	-0.0052	-0.0197	-0.0152	0.0487	0.0682	61.7	70.7
	(0.021)	(0.021)	(0.010)	(0.010)	(0.031)	(0.031)	(0.035)	(0.043)	(52.30)	(53.20)
Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	851	851	787	793	783	793	789	793	789	793
R-squared	0.050	0.056	0.023	0.019	0.838	0.845	0.015	0.022	0.188	0.178

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.

A Appendix

We provide a simple framework to characterize the optimal committee decisions. This framework takes into account the specific institutional features of BancaMia. In particular we assume that the main decision variables of the committee are loan amount and effort to learn about the quality of an application (for simplicity, we ignore the choice of loan maturity).

These simple assumptions are strongly confirmed in the data since we see that committees' main role is to decide loan amounts. In addition, the assumption that scores and committee effort improve the precision of information delivers predictions that are consistent with the fact that treatment effects on loan sizes are mean zero.

We make an additional simplification: the incentive problem of the loan officer is modeled in reduced form. There is an incentive problem in the background that the bank must solve with a combination of formal and informal incentives. Providing incentives to the loan officer costs λ per unit of effort, and by a revelation principle argument, after being adequately incentivized, the officer's effort level is observed.

A.1 Committee Decision Making

The committee receives an application and must choose a loan size, L, and an effort level by the officer, e. The loan produces a gross return to the bank of R with probability $1-p$ and the cost of capital is zero. Default probability is increasing in the amount of the loan and increasing and convex in a borrower-specific risk parameter θ_i:

$$p = L\theta_i^2$$

with $\theta_i << 1$ so that the probability is between 0 and 1.

The committee observes an unbiased (mean θ_i) signal of the risk parameter $\tilde{\theta}_i$, with variance $\tilde{\sigma}_i^2$. The variance of the signal increases with a borrower-specific baseline variance σ_i^2 (reflects how difficult it is to evaluate a borrower's riskiness), and decreasing with the effort of the officer:

$$\tilde{\sigma}_i^2 = \frac{\sigma_i^2}{e}$$

The risk-neutral committee maximizes the expected returns for each loan, net of the cost of incentivizing the officer: $E\left[L\left(1-p_i\right)R - L - \lambda e\right]$. The expectation of the default probability can be re-written as: $E\left[p_i\right] = LE\left[\tilde{\theta}_i^2\right] = L\left[\theta_i^2 + \tilde{\sigma}_i^2\right]$.

Note that the convexity of p on L is a simple reduced-form assumption meant to capture that committee members care about making mistakes in their evaluation of the borrower's risk. This could come, for example, from career concerns: the committee members have a reputation for being able to read a borrower's risk correctly and making mistakes affects their future prospects in the bank. This will make members of the committee risk averse towards the uncertainty on p.

Substituting in the objective function, the expected return from a loan is defined by the following maximization problem:

$E(\pi) = \max_{L,e} L\left(1 - L\left[\theta_i^2 + \frac{\sigma_i^2}{e}\right]\right) R - L - \lambda e$

$s.t.\;\; L \geq 0;\;\; e \geq 0$

The committee observes the borrower-specific parameters θ_i and σ_i^2. The gross return R is also treated as a parameter, since the committee does not decide the interest rate of the loan.

Finally we model the possibility of referrals: the committee can ask for the manager's help at a cost C. When the committee refers the problem to the boss, the boss sees the true risk parameter θ_i, so the net return from a loan after a referral is: $E\left(\pi^R\right) = \max_L L\left(1 - L\theta_i^2\right) R - L - C$.

A.2 In Committee Decisions: Optimal L and e

We first solve for the optimal decision of the committee, conditional on the committee making a decision. Then we explore the decision to refer the loan by comparing the expected return of making the decision in the committee versus referring the application to the boss.

We assume that the parameters are such that we have an interior solution ($L > 0$). The F.O.C. for L is $L(e) = \frac{R-1}{2R\left(\theta_i^2 + \frac{\sigma_i^2}{e}\right)}$, and for e: $e(L) = L\sqrt{\frac{R\sigma_i^2}{\lambda}}$.

Solving gives:

$$L^* = \frac{R-1}{2R\theta_i^2} - \frac{\sigma_i}{\theta_i^2}\sqrt{\frac{\lambda}{R}}$$

$$e^* = \frac{\sigma_i}{\theta_i^2}\left[\frac{R-1}{2\sqrt{\lambda R}} - \sigma_i\right]$$

Note: to have $L > 0$ and $e > 0$ requires that the following condition holds: $\frac{R-1}{2\sqrt{R}} > \sigma_i\sqrt{\lambda}$. This implies that the committee will reject applications that are too difficult (high σ_i) for any given cost of inducing effort.

A.3 Comparative statics

The randomized trial moves exogenously σ_i and/or λ, depending on whether the information channel or the incentive channel are at work. So comparative statics with respect to variations in these parameters give us the predicted changes in actions by the committee in the treatment group.

The optimal loan size is decreasing in σ_i, and λ:

$\frac{\partial L^*}{\partial \sigma_i} = -\frac{1}{\theta_i^2}\sqrt{\frac{\lambda}{R}} < 0$

$\frac{\partial L^*}{\partial \lambda} = -\frac{1}{2}\frac{\sigma_i}{\theta_i^2}\sqrt{\frac{1}{\lambda R}} < 0$

Thus, the introduction of a technology that lowers σ_i or λ, should increase average loan size, ceteris paribus.

Optimal effort is decreasing in λ, but the effect of changing σ_i is ambiguous:
$\frac{\partial e^*}{\partial \lambda} = -\frac{1}{2}\frac{\sigma_i}{\theta_i^2}\frac{R-1}{2\sqrt{R}}\lambda^{-3/2} < 0$
$\frac{\partial e^*}{\partial \sigma_i} = \frac{1}{\theta_i^2}\left[\frac{R-1}{2\sqrt{\lambda R}} - 2\sigma_i\right] \geq 0$ if $\frac{R-1}{2\sqrt{R}} > 2\sigma_i\sqrt{\lambda}$
so $\frac{\partial e^*}{\partial \sigma_i} \geq 0$ for easier applications (applications with a low σ_i) and negative for difficult applications. The effect of a technology that lowers σ_i on average effort will thus depend on the distribution of σ_i in the cross-section of applications.

We cannot take these comparative statics straight to the data because treatment induces changes in two extensive margins : 1) the probability or rejection, and 2) the probability that the committee refers the application to the boss. These comparative statics are based on looking at the same application, before and after a change in the parameter.

A.4 Referrals

The committee refers the application when $E\left(\pi^R\right) \geq E\left(\pi^*\right)$. The optimal loan size with a referral is $L^R = \frac{R-1}{2R\theta_i^2}$. This implies that for any given borrower risk parameter, the boss approves a larger amount than the committee.

A decline in σ_i and/or λ leads to a decrease (weakly) in the number of referrals. This follows from the fact that $\frac{dE(\pi^*)}{d\sigma_i} < 0$ and $\frac{dE(\pi^*)}{d\lambda} < 0$, while $\frac{dE\left(\pi^R\right)}{d\sigma_i} = \frac{dE\left(\pi^R\right)}{d\lambda} = 0$. $\frac{dE(\pi^*)}{d\sigma_i} < 0$ also suggests that the marginal applications, those that are decided by the boss in the absence of scores and by the committee when the score is available, are those that are more difficult to evaluate (although this is a cross-sectional statement and thus depends on the sign of $Cov\left(\theta_i, \sigma_i^2\right)$).

The loan amount drops for the marginal applications because $L^R > L^*$ for all θ_i. This implies that the average effect on the size of approved applications is ambiguous (including those approved by the boss). Ambiguity in the average loan amount also implies ambiguity in the expected default probability.

To characterize the effect on loan size and effort of applications approved by the committee only, one would need to characterize the marginal applications (the least profitable applications referred to the boss). To characterize the marginal applications in the cross-section one needs to know what the joint distribution of θ_i and σ_i^2 is in the data. If we assume they are independent, then $\frac{dE(\pi^*)}{d\sigma_i} < 0$ indicates that the marginal applications will be more difficult to evaluate (have a higher σ_i^2) than the inframarginal ones. This implies that marginal applications require more effort than inframarginal ones.

A.5 Loan Allocation in the Cross-Section

Although the introduction of scores improves the allocation of credit across borrowers, the model has ambiguous implications for the average *level* of observable outcomes (committee effort, loan amount, and default). In order to derive testable implications we are interested in characterizing the cross-sectional implications of improving the efficiency of loan allocation. To evaluate cross-sectional relationships, however, we cannot rely on comparative statics on the parameters, which assume that the parameters are independently distributed in the cross-section of loans. In the data, it is more likely that θ_i and σ_i are jointly distributed with some correlation. We can use the data together with the model to tell us what the correlation between θ_i and σ_i in the population is. This requires taking the model seriously to interpret the relationships observed in the data and make inferences that are not causal but solely identified through the assumptions in the model.

We find a negative relationship in the cross-section between loan amounts and the default probability. Let's start by assuming θ_i and σ_i are independent. If so, then we have that $p = L^*\theta_i^2$, and $L^*\left(\theta_i, \sigma_i^2\right) \sim 1/\theta_i^2$, so p should be constant in the cross-section when there is heterogeneity in both θ_i or σ_i^2. The second one is obvious, since the only effect of σ_i on p is through L^*. The first one can be shown by looking at the total derivative of p with respect to θ:[26]

$$\frac{dp}{d\theta} = \theta^2 . \frac{\partial L^*}{\partial \theta} + 2\theta L^* = \theta^2 \left(-2\right) \left[\frac{R-1}{2R\theta_i^3} - \frac{\sigma_i}{\theta_i^3}\sqrt{\frac{\lambda}{R}}\right] + 2\theta \left[\frac{R-1}{2R\theta_i^2} - \frac{\sigma_i}{\theta_i^2}\sqrt{\frac{\lambda}{R}}\right] = 0$$

Intuitively, to get the negative relationship between default and loan size observed in the data through heterogeneity in θ_i, one requires that L^* drops at a rate higher than $1/\theta_i^2$. From inspection of the equation for L^*, this occurs if θ_i and σ_i are positively associated in the cross-section ($Cov\left(\theta_i, \sigma_i\right) > 0$). That is, if borrowers that are riskier are also more difficult to evaluate. To see this, suppose that the relationship is deterministic and linear: $\sigma_i = a\theta_i$, with $a > 0$. Then:

$$\frac{dp}{d\theta} = \theta^2 \left[\left(-2\right) \frac{R-1}{2R\theta_i^3} + \frac{a}{\theta_i^2}\sqrt{\frac{\lambda}{R}}\right] + 2\theta \left[\frac{R-1}{2R\theta_i^2} - \frac{a}{\theta_i^2}\sqrt{\frac{\lambda}{R}}\right] = -a\sqrt{\frac{\lambda}{R}}\left(\frac{2}{\theta} - 1\right) < 0$$

Note that we can think of the technology that reduces σ_i as a technology that reduces a. So anything that reduces a or that reduces λ will lead to a decline in the absolute value of $\frac{dp}{d\theta}$, or a flattening of the relationship between the default probability and loan amounts in the cross-section. This implication can be directly tested in the data.

[26] In a more direct way, we can evaluate the total derivative: $\frac{dp}{dL^*} = \frac{\partial p}{\partial \theta_i^2}\frac{\partial (L^*(\theta))^{-1}}{\partial L^*} + \frac{\partial p}{\partial L^*} = L^*\left[-\frac{1}{L^{*2}}\left(\frac{R-1}{2R} - \sigma_i\sqrt{\frac{\lambda}{R}}\right)\right] + \theta_i^2 = 0$ where $\left(L^*\left(\theta\right)\right)^{-1}$ is the inverse of the L^* function (θ expressed as a function of L^*).

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 a 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 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) are estimated on the subsample of issued loans. Robust standard errors in parenthesis. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table A.3: Default Probability Correlates in the Cross Section of Control Group Applications

Dependent Variable:	In Default after			
	6 months		12 months	
	(1)	(2)	(3)	(4)
ln(Application Amount)	-0.0069	0.0520	-0.0572**	0.1463*
	(0.012)	(0.049)	(0.029)	(0.079)
ln(Application Maturity)	0.0064	-0.0527	0.0646*	-0.0539
	(0.015)	(0.039)	(0.038)	(0.063)
Credit Risk Score	0.3325*	0.2850*	0.7426**	0.5823*
	(0.181)	(0.164)	(0.325)	(0.320)
First Application	0.0284	0.0208	-0.0171	-0.0329
	(0.041)	(0.043)	(0.052)	(0.055)
ln(Loan Amount)		-0.0777		-0.2545***
		(0.054)		(0.085)
ln(Loan Maturity)		0.0794		0.1563**
		(0.052)		(0.071)
Trend	Yes	Yes	Yes	Yes
Observations	248	248	248	248
R-squared	0.040	0.056	0.058	0.103

Linear probability model of default using ex ante application characteristics and ex post loan characteristics as independent variables. Estimated on the subsample of applications in the control group that were approved by the committee.

Table A.4: 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 decisions 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.

Table A.5: Information versus Incentives: Effect on Interim and Final Actions in $T2$ – ML

Estimation	Interim Outcomes				Final Outcomes			
Action:	Approve (Omitted)	Reject	More Information	Send to Manager	Approve (Omitted)	Reject	More Information	Send to Manager
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment T2		1.5208	-0.4868	-0.8152**		1.4762	-0.5112	-1.5197***
		(1.071)	(0.347)	(0.392)		(1.058)	(0.346)	(0.483)
ln(Application Amount)		-0.1783	0.8793**	-0.1355		-0.1558	0.9058**	0.5286
		(0.698)	(0.359)	(0.453)		(0.701)	(0.358)	(0.501)
ln(Application Maturity)		0.1960	0.0517	1.8806***		0.1588	0.0159	1.0678*
		(1.352)	(0.575)	(0.652)		(1.342)	(0.568)	(0.618)
Credit Risk Score		6.3225**	1.1175	1.7954		6.4498**	1.3050	4.2580
		(2.507)	(1.738)	(2.708)		(2.566)	(1.774)	(2.806)
First Application		0.2935	-0.9911	-0.5589		0.3135	-0.9733	-0.4016
		(0.642)	(0.644)	(0.656)		(0.644)	(0.643)	(0.766)
Trend		Yes	Yes	Yes		Yes	Yes	Yes
Observations	850				850			
R-squared								
Pseudo R-squared	0.0975				0.114			
Fraction in Control Subsample	0.8866	0.0030	0.0627	0.0478	0.8866	0.0030	0.0627	0.0478
Marginal Effects:								
Treatment T2	0.0279	0.0140	-0.0172	-0.0247*	0.0392*	0.0138	-0.0170	-0.0360***
	(0.0201)	(0.0104)	(0.0131)	(0.0128)	(0.0204)	(0.0102)	(0.0129)	(0.0136)

Multinomial Logistic Regression estimates of the effect of treatment on interim committee decisions made before observing the score (columns 1 through 4) and on final decisions made 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.6: Amount Changes: Application, Interim and Final in Treatment $T2$

Comparison:	Interim versus Application Amount (1)	Final versus Interim Amount (2)	Mean Difference (3)
Amount Change Dummy	0.818	0.228	0.590***
Amount After/Amount Before	0.958	1.001	-0.0437***
	(0.305)	(0.107)	(0.349)
|Amount change|	279.0	35.9	243.1***
	(517.7)	(102.5)	(514.5)

Statistics of the frequency and magnitude of loan amount changes that occur between the application and the interim decision (column 1), and also between the interim and the final decision (column 2). Column 3 shows the difference in means. *** indicates significance at the 1% level in a difference in means t-test.