Are the Monitors Over-Monitored? Evidence from Corruption, Vigilance, and Lending in Indian Banks

Abhijit Banerjee, Shawn Cole, Esther Duflo*

April 2008: Preliminary and Incomplete†

Abstract

We analyze a model of lending in which loan officers face both an incentive to lend, and the possibility of penalties for making loans that go bad. The model predicts the following: new firms will face credit constraints; if the punishment does not increase proportionally with the volume of default, loan officers will have an incentive to "evergreen," i.e. bail out failed borrower when the probability of being punished for default increases. We test these predictions using a dataset of all commercial bank loans in India from 1981 to 2003, combined with a data set on the investigation and discovery of fraud by the "central vigilance commission", a federal body charged with investigating default. We find evidence that following the discovery of a fraud in a particular bank branch, vigilance activities greatly increases. This in turn results in reduced lending: the amount of credit declines sharply at the affected bank branch, as well as neighboring branches. This effect is large, and persists in part for up to two years. Bank risk-taking also declines following a vigilance event.

*MIT (banerjee@mit.edu), Harvard Business School (scole@hbs.edu), and MIT (eduflo@mit.edu), respectively.

†We thank Raghuram Rajan, with whom a conversation many years ago prompted us to undertake this project. This work is a collaborative exercise with research staff from the Reserve Bank of India, including S. Majumdar, Deepak Gogia, and Janak Raj. The RBI performed much of the statistical analysis in this paper. The statements made in this paper do not in any way represent the view of the Reserve Bank of India or its employees. We thank seminar participants at Harvard Business School, NEUDC, MIT-Sloan, and Florida Atlantic University. Please do not circulate or cite without authors' permission.

1 Introduction

Standard neo-classical theory predicts that interest rates alone should predict the supply of bank credit, and that the expected risks and returns of various projects and borrowers should determine its allocation. Because of the need to spread idiosyncratic risk, bank have, however, a tendency to be large organizations, which makes providing incentives to loan officers to collect and reveal information about the borrowers inherently difficult. How banks solve these incentives problems can have fundamental consequences on well capital is allocated in the economy.

Several studies have shown that organizational issues within banks do indeed affect the volume and the nature of lending. Bergen, Demsetz and Strahan (1999) show that after consolidations, banks are less likely to lend to small businesses and Stein (2002) argues that this phenomenon is due to the difficulty of providing incentives to gather and truthfully reveal "soft" information within large banks. Hertzberg, Liberti and Paravisini (2007) further investigate the role of incentives to provide information within banks. They show that rotation of officers alleviate the temptation by loan officers not to reveal bad news about the firms they deal with.

The fundamental problem that banks face is that they need to provide incentives to their employees, who are paid a fraction of the money they have to manage, both to exert effort to find clients to lend to, and to avoid default. These two objectives are, however likely to conflict since, after all, the easiest way to avoid a default is not to lend at all.

In this paper, we study, theoretically and empirically, the impact of how loan officers are punished for defaulting on the volume and the nature of their lending decisions. Incentives to prevent default are necessary to induce loan officers to lend to profitable projects, to monitor their clients, and to refrain from stealing money. But they may also discourage honest loan officers to lend. And if the punishment does not sharply increase with the size of the default, an increase in the probability to be punished may also encourage "evergreening", or "gambling for revival": lending to firms which are about to default in the hope that they will receive a good draw in the next period and get out of their bad situation.[1]

Our empirical setting is the public banking sector in India, which has long dominated the banking sector in this country. Public banks have limited availability to provide incentives

[1] Rajan (1994) develop a model where these types of behavior is undertaken by loan officer to protect their reputation.

to loan officers: wages increase mainly with seniority, and it is difficult to fire employees. The main instrument for the bank to punish a loan officer is to transfer him to an undesirable posting. However, the risk of corruption is a key concern to the government and, to alleviate it, loans officers are subject to the government anti-corruption loans. In particular, a bad loan may in principle be treated as embezzlement of government funds, and thus lead to a criminal investigation. A federal body, the Central Vigilance Commission, investigates a small fraction of suspected frauds. While investigations are rare, and convictions even rarer, bankers argue that the risk of being charged with corruption is a "sword of Damocles" hanging above their heads and preventing them to make reasonable decisions.

In this context, we study what happens when the probability of being investigated by the vigilance commission increases. To do so, we exploit an unusual data set, made available at the Reserve Bank of India. The data set contains information (date, type, location, and punishment) on all the frauds reported to the Reserve Bank of India between 1980 and 5006. We use it in conjunction with quarterly and annual information on lending at the branch level for a panel of over 43,000 bank branches in the country followed for several years.

We start by showing that when a fraud is discovered by the Central Vigilance Commission in a particular branch in a particular year, the discovery of other frauds in the same branch increases by a factor of 27 in that year, compared to previous years. This suggests a clear persistence in vigilance activity: the probability to be punished for default is larger in a bank branch in the year following the investigation of a fraud. We exploit this within an event-study approach, where we study how lending behavior in a particular branch changes after the discovery of a fraud in the years following the discovery, compared to other branches. We find a rapid and persistent drop in credit at the affected branch, which cumulates to at least a 20% decline in lending. This decline is driven by a decline in high-quality loans, and we find suggestive evidence of ever-greening. Vigilance activity also affects the sectoral allocation of credit, as branches shift credit from riskier to less risky sectors. Finally, vigilance activity affects other branches of the same bank in the same town.

These results echo those in Banerjee and Duflo (2008), which exploit a shock to the availability of subsidized credit to a subset of firms in India, and find evidence that even fairly large firms with existing relationship to the banking sector are severely credit constrained. These results also were indicating of misallocation of credit by loan officers: the average rupee lent has a much lower return than the extra credit induced by the program; there seem to be no relationship between performance and increment in the lending limit. By showing that corruption has an indirect negative effect on the volume and the allocation of credit in an economy, through its effect on the honest agents who are worried of being punished, they also help explain the large impact of corruption on lending, found for example in Beck, Demirguc-Kunt and Levine (2006).

The remainder of this paper proceeds as follows. In section 2 we provide a brief description of banking in India, and a discussion of the debate on the relationship between corruption and lending. We develop a model of the loan allocation decision by a honest agent subject to incentives to lend and to avoid default in Section 3, and describe the data in the following section. Our first analysis is conducted at the bank-level, in section **??**. We then focus on micro-data, studying where and when frauds are most likely to occur, in section **??**. Section 6 describes the methodology and presents the baseline results. Sections 6.1.2 measures how the effect of discovering credit may spill over to other branches, and 6.2 examines how the nature of lending decisions change after a vigilance inquiry, including the quality of loans, and the risks taken by bank employees. Finally, section 7 concludes.

2 Context & Overview

In this section, we briefly describe the context in which banks operate, including anecdotal evidence on the claim that fear of prosecution causes loan officers to behave very conservatively.

We focus on government-owned banks because they dominate the Indian financial system. Public sector banks include both government founded banks (such as the Bank of India), as well as banks nationalized by the government in 1969 and 1980. In the 1990s, several formerly public banks were partially privatized, and foreign banks had entered, primarily in large cities. While public sector banks market share has declined since then, in 2006 government-owned banks provided 71% of bank credit in the economy.

Corruption in lending is a major policy concern, in India and elsewhere. Cole (2008) shows that lending follows a political cycle. Khwaja and Mian (2005) document political corruption of loans in Pakistan. In India, since public sector banks are owned by the government, employees of the bank are treated by law as public servants, and thus subject to government anti-corruption rules. In particular, any default is in principle subject to investigation by the Central Vigilance Commission, a federal authority which maintains an arms-length relationship with the banks, but is empowered to prosecute loan officers.

Bankers have expressed concern that it is very easy to be charged with corruption. Some feel that any financial loss to a government owned bank would automatically lead to investigation, with the burden of proof on the banker to prove her or his innocence. The quote below is fairly typical.

> [Bankers] say that government officials, who are clueless about banking economics, are posted as vigilance officials and there is the threat of a vigilance inquiry each time a bona fide lending decision results in a bad account. Bankers say that they are flush with funds, but credit is not taking off because of fear psychosis. A top banker said: "You become averse to risks in such an environment. Many of us are turning to narrow banking and stick to gilts.[2]"

The Reserve Bank of India (RBI), the Indian equivalent of the Federal Reserve has also argued that another result of this fear is bankers reluctance to settle non-performing loans (RBI 2000, p. 29), for fear of being charged with corruption. In 2000, the Indian banking system had among the highest percentage of assets in default of any banking system in the world.

A working group on banking policy set up by the Reserve Bank of India, and chaired by M.S. Verma, noted in 2000:

> The [working group] observed that it has received representations from the managements and the unions of the banks complaining about the diffidence in taking credit decisions with which the banks are beset at present. This is due to investigations by outside agencies on the accountability of staff in respect of some of the N[on] P[erforming] A[ssets]. The group also noticed a marked reluctance at various levels to take any credit decision. (Tannan 2001, p. 1579).

[2]Frontline (Hindu), 15(10), May 9-22, 1998. Accessed at http://www.hinduonnet.com/fline/fl1510/15101090.htm on October 20, 2007.

In response to criticism from bankers, economists, and others, the Central Vigilance Commission introduced in 1999 a special chapter of the vigilance manual, on vigilance in public sector banks. While this new chapter was meant to reassure bankers, the language would probably not reassure anyone with experience working in a multinational bank. The manual reads, for example, that "every loss caused to the organization, either in pecuniary or non-pecuniary terms, need not necessarily become the subject matter of a vigilance inquiry... once a vigilance angle is evident, it becomes necessary to determine through an impartial investigation as to what went wrong and who is accountable for the same." (p. 5)

Interviews with public sector bankers revealed widespread concern: the legal proceedings surrounding charges of corruption can drag on for years, leaving individuals charged with corruption in an uncertain state. Even if an individual is exonerated, she may have been relieved of her duties, transferred, or passed over for promotion during the time of investigation. In theory (as well as practice), even one loan gone bad may be sufficient to start vigilance proceedings. The possible penalties stand in stark contrast to rewards. While banks are constantly urged by the Reserve Bank of India to loan as much as possible, there are no explicit incentives for making good loans, or ways to penalize officers who make conservative decisions.

Not surprisingly, the Central Vigilance Commission disputes the claim that there is a "fear psychoses," and, to bolster their position, released in 2000 a "critical analysis" of vigilance activity in public sector banks in 1999. The analysis reveals that in 1999, the Central Vigilance Commission received 1,916 references, 72% of which were credit-related, of which 55% resulted in recommendations for major punishment.

Thus, compared to the volume of lending, the probability to be investigated for a fraud is actually extremely small. Moreover, their 2000 report states "Out of every 100 cases coming before it, the Commission would advice major penalty proceedings in 28 cases, minor penalty proceedings in 32 cases, and administrative warning/exoneration in 40 cases." (p. 9). The author of the report, a CVC official, argued that this level of activity should not be enough to cause "fear psychoses": "These figures reveal that a person is not damned the moment his [sic] case is referred to the Commission...These statistics appear to indicate a very fair and objective approach on the part of the Commission to the cases that were referred to it." (CVC 2000 p. 10).

Nevertheless, the "fear of lending" arguments made by the bankers seem to have been serious enough to convince the government: In April 2004 (after the period covered by this study), lower-level loan officers were removed from the jurisdiction of the Central Vigilance Commission. Our data covers the period before this changes, when even low level loan officers could still be charged by the CVC.

3 Theory

In this section, we describe a model in which a loan officer allocates capital to firms. Loan officers are given incentives to lend money and they are punished for default. The model studies how the shape of the punishment as a function of the default affect a honest loan officer's incentive to lend, and to allocate its lending.

3.1 Model setup

Firms: Firms in this economy come in two types–H and L–in proportions p and $1-p$. Their production function is as follows: with probability $p(k,\theta_i)$, where $i = L, H$, $\theta_H > \theta_L$ and $p(k,\theta_i)$ is a concave function of k, the firm produces μk. With probability $1-$ $p(k,\theta_i)$ the firm produces 0. Assume that $p_k(k,\theta_i) < 0$, i.e. bigger projects are more likely to fail, that $p(k,\theta_H) > p(k,\theta_L)$ which defines the type H firm to be the more productive firm, and that $0 > p_k(k,\theta_H) > p_k(k,\theta_L)$. That is, p_k is increasing in θ, which derives from the idea that the better firm is better able to handle an inflow of capital. Finally assume that $kp(k,\theta_i)$ is an increasing but concave function of k.

Firms live for two periods and produce in both–the shock to its productivity is drawn independently from the same distribution in both periods. Firms do not save and have no equity, so in each period, production is entirely financed by credit and the current output is all they have to repay any loans. We will assume that there is no firm-side moral hazard: firms never deliberately default. However if they do not have enough money they have no choice but to default.

Bank: There is a bank that is the sole source of capital to this population of firms. To simplify matters we will always assume that banks set a single interest rate r for all loans. The cost of bank capital is $\rho < r$. Lending decisions are however made by bankers who work for the bank but maximize their own expected earnings (rather than the bank's earnings), which in turn are a function of the incentives that the banks set for them.

Bankers: Each new firm get allocated to a banker who work with it for the next two periods. Bankers get paid an amount C per unit they lend. They also get punished for any loans that they made that were ultimately defaulted upon. The punishment is given by $\tau F(k)$ where k is size of the loan, F is an increasing function and τ is a shift parameter which distinguishes between high and low monitoring regimes. The fact that both C and $F(k)$ are taken as being exogenous is unorthodox. Endogenizing C is relatively straightforward. As far as endogenizing F is concerned, the fact is that it is determined by the Vigilance Commission, which deliberately maintains an arms-length relation with the bank. Therefore we should not expect it to be chosen optimally from the point of the bank. Our strategy therefore is to treat it as a parameter and examine the effects of shifts in it. Finally we assume that bankers do not discount the future and maximize expected net benefits (i.e. are risk neutral).

Information: In the first period neither the firm nor the banker knows the firm's type. At the end of the first period they both observe the firm's output and update accordingly using the fact they both know how much the firm had invested: Successful firms are going to be type H with probability

$$p_s(k) = \frac{pp(k, \theta_H)}{pp(k, \theta_H) + (1 - p)p(k, \theta_L)}$$

Since $0 > p_k(k, \theta_H) > p_k(k, \theta_L)$ and $p(k, \theta_H) > p(k, \theta_L)$, $\frac{p(k,\theta_H)}{p(k,\theta_L)}$ is increasing in k, which is why p_s is increasing in k. Likewise the probability that a failed firm is type H, p_f, is declining in k.

Timing: We will study the case where the current level of monitoring (and hence the expected punishment) is high but is expected to go down. In particular assume that a present level of $\tau > 1$, but that it is expected to go down to $\tau = 1$ over the next periods. In particular, the probability that it goes down to 1 in the next period is q and if it does not go down next period it is certain to go down in the following period. Once it does go down it is expected to stay there for the foreseeable future. The idea is that the bank is now under scrutiny so that defaults are highly likely to be investigated, but eventually the vigilance commission's attention will shift elsewhere and the probability of an investigation will go down.

A banker starts with one population of borrowers who borrowed last period (before τ

went up) and another population of new borrowers.

3.1.1 Lending to a successful old firm

Suppose that a firm is in the second period of its life, was able to repay its first period loan and is of type H with probability p'. Then the amount the banker would want to lend to that firm in the second period is given by maximizing

$$Ck - [p'(1 - p(k_2, \theta_H)) + (1 - p')(1 - p(k_2, \theta_L))](q + (1 - q)\tau)F(k)$$

where we write k_2 to remind ourselves that this is the second period. The solution is given by solving the equation

$$\begin{aligned} C &= [p'(1 - p(k_2, \theta_H)) + (1 - p')(1 - p(k_2, \theta_L))](q + (1 - q)\tau)F'(k_2) \\ &\quad -[p'p_k(k_2, \theta_H) + (1 - p')p_k(k_2, \theta_L)](q + (1 - q)\tau)F(k_2) \end{aligned}$$

As long as the second order condition for this maximization holds (which we assume), an increase in p' will increase k_2, because both $[p'(1 - p(k_2, \theta_H)) + (1 - p')(1 - p(k_2, \theta_L))]$ and $-[p'p_k(k_2, \theta_H) + (1 - p')p_k(k_2, \theta_L)]$ go down when p' goes up. Better borrowers get bigger loans. Also k_2 is declining in $q + (1 - q)\tau$, which is the expected value of τ, which we will denote by τ^e. More stringent punishments will discourage lending. Hence $k_2 = k_2(p', \tau^e)$, with $k_{2p} > 0$, and $k_{2\tau^e} < 0$.

Clearly, for appropriately chosen values of F and C, the amount the banker wants to lend can be very large, indeed larger than what the firm wants to borrow. However let us assume that the opposite is the case in the empirically relevant range—which is consistent with the evidence in Banerjee and Duflo (2004) showing that clients of the banking system tend to be credit constrained.

3.2 Lending to a failed old firm: bailouts

Consider now the alternative case where everything is the same about the firm except that it has failed in the first period and hence has no output to repay its loan. One thing the banker could do is to let the firm declare bankruptcy and then lend to it based on its p' and the value of τ expected for the next period, i.e. $k_2(p', \tau^e)$. In this case the banker gets a punishment $F(k_1)$ for the first period default, where k_1 was the first period loan to the firm. The alternative is to bail it out: Give it a loan such that it can repay the first period loan (and thereby avoid defaulting) and as well as the second period loan, if it manages to the successful in period 2. This means a loan of size $k_b(k_1)$

$$\begin{aligned}\mu(k_b(k_1) - k_1 r) &= k_b(k_1) r \\ \text{or } k_b(k_1) &= \frac{\mu r k_1}{\mu - r}.\end{aligned}$$

This would avoid immediate default but create a risk of a bigger default in the future (since $r > 1$, $\frac{\mu r}{\mu - r} > 1$). From the banker's point of view the trade-off is between

$$\begin{aligned}&\Pi_{NB}(k_1, p', \tau^e, \tau) \\ =\ & Ck_2(p', \tau^e) - \tau F(k_1) - \tau^e F(k_2(p', \tau^e)[1 - p' p(k_2(p', \tau^e), \theta_H)) - (1 - p')p(k_2(p', \tau^e), \theta_L))]\end{aligned}$$

and

$$\begin{aligned}&\Pi_B(k_1, p', \tau^e) \\ =\ & Ck_b(k_1) - \tau^e F(k_b(k_1)[1 - p' p(k_b(k_1) - rk_1, \theta_H) - (1 - p')p(k_b(k_1) - rk_1, \theta_L)]\end{aligned}$$

Denoting by $1 - p' p(k, \theta_H)) - (1 - p')p(k, \theta_L) = \pi(k, p')$, bailouts dominate if

$$\begin{aligned}&\tau F(k_1) \\ \geq\ & \{Ck_2(p', \tau^e) - \tau^e F(k_2(p', \tau^e))\pi(k_2(p', \tau^e), p')\} \\ & -\{Ck_b(k_1) - \tau^e F(k_b(k_1))\pi(k_b(k_1) - rk_1), p'\}\end{aligned} \tag{1}$$

It can be shown that $k_2(p', \tau^e) < k_1$ and hence $k_2(p', \tau^e) < k_b(k_1)$. Bailouts are costly because they lead to over-lending. The cost of this distortion is what the term on the right-hand-side represents: The first term is payoff from what the banker would ideally want to lend, while the second is the payoff from what he ends up having to lend in order to bailout the firm. The benefit is the term on the left: he avoids paying the penalty on an immediate and certain default.

To understand the factors that make bailouts more attractive, first consider two extreme cases: One in which F is very convex, so that both $F(k_1)$ and $F(k_2(p', \tau^e))$ are small relative to

$F(k_b(k_1))$ (the extreme case of this is when $F(k_1) = F(k_2(p', \tau^e)) = 0$, but $F(k_b(k_1))$ is large). In this case, bailouts will not typically be worthwhile.

In the alternative case where F is very concave goes the other way. The extreme case of this is where F is (almost) a constant, so that $F(k_1) \approx F(k_2(p', \tau^e)) \approx F(k_b(k_1))$. In this case bailouts almost always pay off.

Next, consider the effect of an increase in τ. Using the envelope theorem, the effect on the right-hand-side of condition 1 can be shown to be equal to

$$\frac{d\tau^e}{d\tau}[F(k_b(k_1))\pi(k_b(k_1) - rk_1, p') - F(k_2(p', \tau^e))\pi(k_2(p', \tau^e), p')]$$

while the effect on the left-hand-side is simply $F(k_1)$. To see which side dominates, consider the value of τ (assume that there is such a value) where the banker is just indifferent between bailing out and letting the firm default, i.e.

$$\begin{aligned} & \frac{\tau}{\tau^e}F(k_1) + C\frac{[k_b(k_1) - k_2(p', \tau^e)]}{\tau^e} \\ = \quad & F(k_b(k_1))\pi(k_b(k_1) - rk_1, p') - F(k_2(p', \tau^e))\pi(k_2(p', \tau^e), p') \end{aligned} \tag{2}$$

At this value of τ, it is clear that it must be the case that $F(k_b(k_1))\pi(k_b(k_1) - k_1, p') - F(k_2(p', \tau^e))\pi(k_2(p', \tau^e), p') > F(k_1)$. Therefore which side of 1 goes up faster when τ goes up, turns on the size of $\frac{d\tau^e}{d\tau} = 1 - q$. If q is large enough, so that by the time the next default becomes a possibility it is likely enough that the vigilance commission has moved on, then increasing τ will raise the right-hand-side of 1 more than the left-hand-side (starting at the point of indifference) and bailouts will be more likely at high current levels of τ. On the other hand if the probability of an investigation is likely to stay relatively constant for some time, then a higher value of τ will make bailouts more costly and will discourage them.

A temporary increase in the likelihood of an investigation may therefore lead to increased bailouts–especially if the rate of investigation is expected to fall off relatively quickly. In terms of observables this might mean less defaults (at least contemporaneously) followed by an increase in default of older firms in the next period.

The intuition behind this result should be evident from thinking about an extreme case: Say $q = 1$, so that we know that next period default will be punished less. Then higher the value of τ today, the bigger the gain from postponing the default by one period and hence greater the incentive to bailout.

Note that we have so far said nothing about whether bailouts are desirable or not from the point of view of the bank, which comes down to whether

$$\begin{aligned}&(r(1-\pi(k_2(p',\tau^e),p')-\rho)k(p',\tau^e)\\ \geq\ & rk_1+(r(1-\pi(k_b(k_1)-k_1),')-\rho)\frac{\mu r k_1}{\mu-r}\end{aligned}$$

or not. The reason why bailouts are not necessarily always bad is that from the bank's point of view the bankers may be too reluctant to lend, because they are responding to the current level of τF and not necessarily to what the bank wants them to respond to. Bailouts may be a way to get the bankers to lend more. However they are never going to be a particularly good way to loans out of the door, since they target the extra loans to failed (and therefore likely to fail) firms. In particular if high type firms are sufficiently likely to succeed and low type firms are sufficiently likely to fail, bailouts will be inefficient. We will implicitly assume that we in a setting where this is true, which is consistent with the standard view that bailouts are undesirable.

3.3 Lending to a young firm

Consider the problem of lending to a young firm in a period where τ is still high,but expected to go down to 1 either in the next period (with probability q) or, failing that, in the one after. Over these two periods the banker maximizes

$$\begin{aligned}&Ck_1+[Ck_2(p_s(k_1),1)-F(k_2(p_s(k_1),1))\pi(k_2(p_s(k_1),1),p_s)](1-\pi(k_1,p))\\ &+q\pi(k_1,p)\max\{\Pi_{NB}(k_1,p_f,1,1),\Pi_B(k_1,p',1)\}\\ &+(1-q)\pi(k_1,p)\max\{\Pi_{NB}(k_1,p_f,\tau,1),\Pi_B(k_1,p',1)\end{aligned}$$

It is easy to see that the benefit from a loan that is successful in the first period is unaffected by τ. On the other hand the benefit/loss from a loan that goes bad (in the case where τ remains high for the next period) is obviously lower when τ is higher. Since lending less increases the likelihood of the success state, a higher value of τ will discourage lending to young in the first period.

Less lending in the first period has a number of implications. First, it reduces the likelihood that any firm will fail in the second period. Second, and relatedly, being a successful firm becomes less informative as a signal for the quality of the firm. As a result of these two effects, the loan contraction might persist beyond the current period even though the punishment goes down to its historical level after the end of the current period, and therefore loans made in the current period are the last group that faces a high penalty level. Basically, the first effect means that there are less bailouts and since bailouts tend to be large, this might depress lending levels. The second effect tells us that success is less informative and therefore p_s will be lower. As a result successful firms will get smaller loans, though there will be more successful firms.

3.4 Summary

A loan officer faced with a failing firm faces a trade-off: lending today will cover up the bad loan, and with luck, the firm may recover. Revealing the bad loan today may result in a penalty for the loan officer, but she will then be able to make the optimal decision going forward. The model derived above makes a series of predictions about loan officer behavior, which we take to the data in the next section.

These predictions depend on the shape of the penalty function $F(.)$ as well as the transition probability for the risk of punishment (q). In particular, a concave $F()$ function will encourage bailouts more than a convex $F()$. Behavior in this period is also governed by the expectation regarding the probability to be charged in the future. Consider a period in which q is high. If the level of monitoring in the subsequent period is also high (q is small), then bailouts will be less attractive, relative to a situation in which the level of monitoring is likely to be low (q is large).

Below (in section **??**), we argue that, in the application we study, $F(.)$ is likely to be concave, and that, following the discovery of a fraud, the risk of detection of future defaults is also abnormally high. In this case, the model generates the following predictions:

1. Lending will decline when monitoring increases, and the effect may be persistent.

2. Banks will prefer to lend to less risky firms, and this effect (under certain assumptions) is stronger in regimes where monitoring is high

3. Increased monitoring leads to evergreening, which will induce a specific patter in bad loans: an decline in default in the period where monitoring is high, followed by a jump in default

rates in the next period (as some of the firms which benefitted from evergreening in the high-monitoring regime fail). The following period, however, default may again fall, if the credit contraction effect is persistent.

4 Data and Descriptive Analysis

This study takes advantage of an unusually rich collection of datasets, collected by the Reserve Bank of India.

The fraud-level data set contains information on frauds and suspected frauds reported to the Reserve Bank of India. Upon discovery of a fraud or suspected fraud, banks are required to file a report with the Reserve Ban of India. In principle, the dataset should have an entry for every fraud reported; however, in the early 2000s, the dataset was transferred to a new electronic records-keeping system, and not all of the older frauds were transferred. For each fraud, we observe the date the fraud occurred, the date the fraud was detected, the branch at which the fraud occurred, and the size of the fraud. Other optional fields include the outcome of the investigation, what punishment was meted out, and whether criminal charges were filed. Unfortunately, these fields are not mandatory, and are populated for a very small number of frauds. Summary statistics are given in Table 1. In total, we observe 862 credit-related frauds, the vast majority of them committed between 1990 and 2002, the period covered by our data set. Of those, 323 are labelled as "extension of credit for illegal gratification" and most of the others are "other credit related" fraud. The large majority of the frauds take place in the public banks (which is not surprising since most of the credit take place in the public banks).

To start providing a sense of this data, table 2 provides evidence on where frauds tend to be detected. We estimate a cross-sectional linear probability model, with the dependent variable equal to 100 if a fraud was detected in that branch at any point between 1980 and 2005:

$$y_b = \alpha + X_b\beta + \varepsilon_b \tag{3}$$

, where X_b is a set of branch characteristics.

Results are presented in Table 2. Column (1) reports results from equation (3), for all 56,446 branches in the sample. As there were only 898 credit-related frauds observed in that period, the magnitude of the coefficients is low. There is a strong positive association between the size of the city in which the branch is located, as with the size of the bank branch, as measured by the log of total credit outstanding in 1998. Being located in a larger credit market (as measured by loan size) is also positively correlated with the detection of a fraud. Curiously, the correlation between the level of corruption in a state (as measured by Transparency International India) and the discovery of fraud is negative. This may indicate that fraud is more likely to be *detected* in states with lower levels of corruption. In the simple cross-section, there is a weak negative relationship between the quality of intermediation (as measured by the share of nonperforming credit, or the percentage of loans late). Columns (2) and (3) present equation (2) with district and credit-market fixed effect, respectively. Most results stay qualitatively similar.

Bank heterogeneity matters as well: in a simple regression of credit fraud on bank fixed-effects, the bank fixed-effects are significant at the 1 % level. In the analysis below, we systematically control for bank-branch fixed effect, to control for systematic differences between branches were fraud tend to be discovered and other branches.

We combine this fraud database with two databases of bank credit, one at quarterly frequency, one at annual. The quarterly dataset gives the aggregate amount of lending and deposits of each bank branch, in India, from June 1991 to May 2006.[3]. The dataset is an unbalanced panel. Our regressions include 56 quarters, covering approximately 43,000 bank branches, for a total of 2.5 million observations.

The second set of credit data is a loan-level, annual data set. Each year, every bank branch in India is required to provide information on every loan in its portfolio to the Reserve Bank of India. This information includes the size of the loan, interest rate, and performance status, as well as various characteristics of the borrower, including industry (at the three-digit level), rural/urban status, etc.[4] We analyze the aggregated data only (aggregations are described below). At the branch level, annual credit growth is available for 11 years, totaling 408,555

[3]In this work, we include all scheduled commercial banks (public, private, and foreign), but exclude Rural Regional Banks, which are not scheduled commercial banks, and therefore not subject to the same set of regulations.

[4]Banks were allowed to report loans smaller than Rs. 25,000 (ca. $625) in an aggregated fashion until 1999, at which point loans below Rs. 200,000 (ca. $5,000) were reported as aggregates.

observations.

The two datasets are complementary: the higher frequency of the quarterly dataset allows for better measurement: frauds are of course discovered throughout the year, and with quarterly data, we are better able to pin down the short- and medium-term effects of the discovery of fraud. The quarterly data also provide a longer time series, of 15 years. To delve deeper into the effects of vigilance activity, we use the annual dataset, which includes information on the repayment status of loans and the industries to which banks lend.

Both the quarterly and annual datasets identify bank, branch, credit market, district, and date. Approximately 10% of the credit data is missing: this occurs when a bank does not report required information.

Taken together, this data form, to our knowledge, one of the richest data sets on fraud in the banking sector, and its implication for lending. Due to the highly confidential nature of the data, all data analysis was performed at the Reserve Bank of India, either by one of the researchers on site, or using codes that was provided by us.

5 Fraud Discovery and Punishment

5.1 The Impact of the Discovery of a Fraud on Vigilance Activity

To test the main predictions of the model, our empirical strategy will be to exploit the increase in the probability of being investigated or punished for default after a fraud has occurred in a particular branch. In the model, the probability that the regime of high inspection continues into the next period is an important parameter: it makes sense for a loan officer to bail out a firm during a "high inspection regime" if she expects that the probability of inspection will be lower in the future. Thus, this section focuses on how the probability of fraud detection evolves over time following the discovery of a fraud in one particular branch.

As we already noted, the probability of detection is very low: the probability of any given branch reporting a suspected fraud is only 1.2% over the entire 15-year period of our data. In any given year, the unconditional probability of vigilance activity is 8 basis points. However, there is very large serial correlation in inspections. To understand how an investigation affects the subsequent probability of a detection ocurring, we estimate the following model:

$$D_{b,t} = \alpha + \sum_{n=1}^{5} D_{b,t-n} + \varepsilon_b \tag{4}$$

where $D_{b,t}$ is a dummy variable taking the value of 1 if the branch reports a suspected fraud in year t, and 0 otherwise. Regression results are displayed in Panel B, for a lag structure including 1-5 lags. Two results stand out: first, if a fraud was detected in a given year, the probability of detection in the subsequent year, relative to a branch in which no fraud was detected, is substantially increased. The coefficient on $D_{b,t-1}$ is .023, indicating an increased probability of 2.3 percentage points. Relative to the baseline probability of 8 basis points, this represents a an increase by a factor of 29 times. Second, the probability of heightened vigilance drops by half the following year, and remains at that level after two years. After three years, however, the probability of detection is very low, indistinguishable from the probability of detection in a branch with no history of fraud detection.

Part of the increase in the probability of detection certainly comes from the fact that the same set of people are likely to have committed more than one fraud. The investigation of one fraud reveals all the older ones. In fact, this is probably what prompts the vigilance team to investigate this branch thoroughly. But as they do so, it is also likely that they investigate the lending activities of others in the bank. Thus, we take this data to suggest that fraud discovery in a bank creates a large increase in the probability of subsequent investigation, which persist for about two years. In the next section, we will investigate how lending decisions respond to this temporary increase in vigilance activity.

5.2 The Shape of the Penalty Function

The shape of the penalty function is an important parameter in the model. With a $F(.)$ sufficiently convex, ever-greening could be entirely avoided.

While it is difficult to get data that precisely pins down the shape of the penalty function, discussions with bankers and media suggest that it is increasing, but rather flat: the standard penalties for having found guilty of frauds is a reduction in pay, transfers, and in the worst case, dismissal. This holds true in our dataset; visual inspection (a categorization was not available) suggested pay reduction as the primary mechanism.

Further evidence that then penalty function is flat comes from an exercise conducted by the Central Vigilance Commission. In order to stigmatize corruption, the CVC decided to publish on its website the names, and penalties, of senior-level officers who had been found guilty of corruption. While the specifics of the fraud committed were not disclosed, because the bankers worked at the highest level, the frauds likely involved very large sums. Of the 76 bank officials for whom data was made public, the distribution of penalties was:

Outcome	**Frequency**
Criminal Charges Filed	1
Censured	2
Obligatory Retirement	4
Dismissed	21
Pay Reduction	46

This, in our mind, provides strong evidence that the penalty for fraud is not commensurate with the amounts involved, let alone convex: a large part of the cost seems to be the investigation itself, which appears to be quite traumatic for low-level loan officers. But the punishments, even for those large fraut, mainly involved a pay reduction. This of course begs the question of why the penalty function is not made more convex: this could be achieved by not prosecuting loan officer till several of their loans appear to have a suspected fraud, and by handing out much more severe punishment once someone is actually investigated and proven guilty. The temptation would be very severe for bankers accused of large fraud to bribe the CVC investigators. The necessity to maintain an arm's length relationship between the bank administration and the CVC may be the reason why the penalty function cannot be convex.

6 Empirical Tests of the Model

6.1 Lending Decisions

We now turn to testing a central prediction of our model: does vigilance activity reduce lending activity? To investigate this question, we treat the discovery of a fraud in a branch as an "event" and study how lending in this bank changes after the discovery of the fraud, compared to other unaffected branches.

6.1.1 Branch-Level Effects

We use an event-study (difference-in-difference) methodology to measure the impact of vigilance activity on lending. This strategy compares how lending evolves in branches that are affected by fraud investigations, to other similar branches that are not.

The base specification uses quarterly lending data, at the branch-level. Denote y_{obct} as the change in log credit at bank branch office o, belonging to bank b, in credit market c, at time t. The estimated equation is thus:

$$y_{obct} = \alpha_o + \sum_{k=-8}^{8} \beta_k D_{o,t,k} + \beta_{\geq 9} D_{o,t,\geq 9} + \gamma_t + \varepsilon_{obct}, \tag{5}$$

where y_{obct} is log credit growth, defined as $\log(credit_{obc,t}/credit_{obc,t-1})$.[5] $D_{o,t,k}$ indicate the *detection* of the fraudulent act. Thus $D_{o,t,0}$ is set to one if a fraud occurs in branch office o in quarter t, and zero otherwise, while $D_{o,t,-3}$ is set to 1 if a fraud will occur in branch office o at time $t+3$. $D_{o,t,\geq 9}$ is dummy variable indicating a fraud was committed nine or more quarters previous to quarter t in branch office o. a_o and γ_t denote branch and quarter fixed-effects, respectively, which we include in the preferred specification.

Table 4 reports results from the baseline specification, for the sample of frauds coded as "extension of credit for illegal gratification" or "other, credit related." Under the null hypothesis that vigilance activity does not affect lending, the coefficients β should be zero.[6] Column (1) reports results from Equation (5) without branch or time fixed effects. Column (2) adds quarter fixed effects (γ_t); column (3) adds branch fixed effects (a_0); and column (4) presents the results with quarter and branch office fixed effects. Column (4) is the preferred specification, as the branch fixed effects control for all unobserved time-invariant heterogeneity, while the quarter fixed effects control for aggregate changes in credit at the national level.

This regression is perhaps best understood graphically: Figure 1 displays the coefficients from the preferred specification for each event time (eight quarters prior to an event, seven quarters prior, etc...) The blue line gives the point estimate, while the yellow lines give the 95-

[5]We use log credit growth, rather than log credit, because credit is highly serially correlated, and because credit growth may be more relevant than the stock of credit.

[6]The average size of a fraudulent loan is a quite small relative to the branch portfolio, so there is no mechanical relationship between discovery of fraud and the amount of lending.

percent confidence interval. The line therefore traces how credit evolves prior to the discovery of a fraud, and following a fraud.

A clear pattern emerges. The corrupt loan is discovered at time zero. Prior to this discovery, credit typically evolves at a rate no different than at other branches, with the exception of "six months prior" and "three months prior," when the growth rate is above average (this may have been what prompted the investigation, or this may be a symptom of fraudulent activity). Following the discovery of fraud, credit growth drops precipitously, relative to other branches: by 1.36 percent relative to branches in which a fraud is not discovered in the quarter of discovery (column (4)). In the subsequent quarter, credit growth is 3.73 log points below unaffected branches. This is a very large effect: the average growth rate in credit is 3.5% per quarter, meaning that following an investigation, the level of credit lend out is entirely stagnant. Credit continues to grow at a slower pace following the discovery of the fraud, with a cumulative effect of nearly twenty percent less credit lent by an affected branch after two years. The final coefficient, $D_{o,t\geq 9}$, is negative and significant, suggesting that credit continues to decline (relative to other branches) even two years after the discovery of the fraud.

Even though our dependent variable is change in log credit, specification 5 is likely to suffer from serial correlation. Running a regression with branch and quarter fixed effects, and 2.5 million observations, correcting for non-independence of error terms in SAS taxes the available computer facility, so the majority of the regressions presented in this paper are currently presented with unadjusted standard errors.[7] However, in an appendix table we present the baseline specification, with corrected and uncorrected standard errors. Correcting the standard errors has little effect on the measured precision of the estimates: only one "pre" coefficient is statistically distinguishable from zero, the "post" coefficients remain negative and similar in size, though three of the eight are not statistically distinguishable from zero. The t-statistic on the effect one quarter after discovery of the fraud remains very high, approximately 5.5.

Table 5 estimates equation 5 with a longer series of post-event dummies, tracing the effect up to two years later. The results for the branch-level identification are presented in column (1). Including sixteen post dummies reduces the sample size by two years of fraud discovery, but provides a longer picture of the effect. The results are nearly identical: the effect

[7] This is due to time constraints. Subsequent versions of this paper will have corrected standard errors.

of discovery of fraud is negative and significant for the first eleven quarters following discovery, with a cumulative effect of 21 percent less credit.

What explains the persistence of the effect on lending, despite the fact that the increase in the vigilance activity in the branch stops relatively quickly? Our model predicts some temporal propagation of shocks. First, mechanically, the volume of loans will be reduced as the vigilance officer discover ever-greening. Second, as loan officers fear increased scrutiny, they reduce the size of loans to new firms. Because the informativeness of a failure is increasing in the size of the loan, loan officers on average learn less about new firms in the vigilance period; then, even when the monitoring threat has reduced, the optimal loan size is smaller. However, firms in our model live only two periods, and the model cannot explain as much persistence. In reality, firms leave longer than two periods, and thus the effect may persist longer. The firms who were kept on alive by successive evergreening, and who were cut from the portfolio during the vigilance period, may in particular have otherwise continued to receive credit for a very long time. It is therefore difficult to conclude whether this persistence is a good thing or a bad thing. Another reason for the stickiness of the effect may be because the investigation remains salient in the minds of loan officers, even as the probability of investigation decreases dramatically.

6.1.2 Spillover Effects of Vigilance Activity

In this section, we investigate whether the investigation has spillover effects to other branches in the same area: do other branches of the same banks in the town also stop lending when a particular branch is investigated following the discovery of a fraud? How about other banks in the same town?

To interpret these results, we first need to investigate whether discovery of fraud in a branch increases the probability a fraud will be discovered in other branches of the same bank in the next year, and other branches in the same town. We investigate this question by by augmenting equation (4) with a dummy for whether a fraud was discovered in another branch of the same bank in the same town in the previous year, and a dummy for whether a fraud was discovered in another branch in the town in the previous year. Both effects are significant, although they are smaller than the direct effect: when a branch of the same bank in the same town, the probability of inspection increases by 0.9 percentage point (about half the size of the direct effect). The increase is only 0.2 percentage point if another branch of the same town, but in a different town, was investigated, although the increase number remains significant.

Since there was an effect on the probability of inspection, we may expect an effect on lending on those branches as well.

These effects can be identified by extending the framework above. First, to simplify notation a little, define the $\phi\,()$ function as a set of dummy variables that indicate an event window for a particular category (e.g., bank branch, or town):

$$\phi\,(Category) = \sum_{k=-8}^{16} \beta_k D_{k,category} + \beta_{\geq 9} D_{\geq 9,category}$$

Thus, $\phi\,$(branch) indicates a series of 25 dummy variables, in the same fashion as the dummy variables introduced above in equation 5. (Here we extend the window from 8 quarters before through sixteen quarters after, along with a dummy for more than 16 quarters following the discovery of a fraud.) Similarly, $\phi\,$(Town) is a set of 25 dummy variables comprising an event window around a fraud being discovered in a particular town: the indicator $D_{c,t,0}$ is 1 for all branches within a town if a fraud was discovered that quarter at any branch within that town. Finally, $\phi\,$(TownBank) gives a set of dummy variables, where the event is discovery of a fraud at any branch office of that particular bank in that town.

To estimate the effect of discovery of fraud at a particular branch on that branch, other branches within its network, and the aggregate supply of credit in the credit market, we estimate the following equation:

$$y_{obct} = \alpha_o + \phi\,(branch) + \phi\,(banktown) + \phi\,(town) + \gamma_t + \varepsilon_{obct} \qquad (6)$$

Table 6 presents regression 6, including 16 lags of after the discovery of the frauds. The seventy-two coefficients of interest (24 for each set of dummies) are presented in columns (1), (2), and (3). The most important effect is clearly at the affected branch, consistent with what we had estimated before: these coefficients are all negative and significant following the discovery of fraud. Consistent with the probability of detection results, there is also and immediate decline in other branches of the same bank as the bank in which the fraud was discovered (column (2)): credit declines by 0.76 percent. The effect is somewhat lower than half of the direct effect on the bank branch (which is what we might have expected if the impact of the probability of detection on lending) but in the the right order of magnitude. This effect is short-lived, however. There does not appear to be a marked increase in lending by other branches in response to the reduced supply of credit. This again is consistent with the much lower impact of lending on those branches. Interestingly, we do not see an *increase* in lending by the other branches, which could have happen if all the clients of the branch that stopped lending moved to other banks.

These results are of independent interest. They are provide also some reassurance that the results are indeed due to the causal effect of the increase in the probability being charged for a fraud, since, from the point of view of other branches in the same town, the discovery of a fraud in another branch is an exogenous event.

6.2 Repayment and Risk-Taking

Section 6 clearly established that credit in a branch (and indeed, a town) declines following the discovery of a fraud. In addition, the model predicts that banks will prefer to lend to safer firms.

To conduct this analysis, we use the loan-level data, which are annual credit data, available from 1992-2003. While these loan data are remarkably comprehensive (recall we have a separate observation for each loan above a relatively low threshold), they are unfortunately not a panel, meaning we cannot link loans from one year to the next. (We can, of course, link the loan to the branch, and link the branch from year to year).

We first replicate the results found at the quarterly level in the annual data, using a window of two years. The results are presented in Table 7. The effects at the branch level are very similar: in the year of a fraud discovery, annual credit growth declines by two percent relative to other bank branches. In the subsequent years, the annual effect is close to six percent, which is roughly consistent with the two-percent quarterly effect. Loan growth prior in branches in which fraud is about to be discovered is no different than growth in branches in which no fraud will be detected. Column (1) presents the results without bank branch fixed effects, while column (2) gives the results with branch fixed effects. Both estimates include year fixed effects.

6.2.1 Branch Risk-Taking

The quantity of credit lent declines rapidly following the discovery of a fraud. What happens to the composition of lending? We first focus on the riskiness of loans made by loan officers:

if lenders are afraid that a bad loan may land them in hot water, they may prefer to lend to borrowers who are less likely to default. We do not observe the credit rating of the borrower (for this time period, there is no formal credit rating institution, and this data set does not include data on the credit-worthiness of the borrower). Instead, we exploit the fact that some industrial sectors are riskier than others. We calculate how risky each industrial sector is, measuring the percentage of loans in that sector that are late in repayment. This measure could vary from zero to 100:

$$sharelate_i = 100 * \frac{\text{Value of Lending in Industry that is Non-Performing in 1992}}{\text{Lending in Industry in 1992}}.$$

We then take the share of credit lent by each branch o to industry I, to come up with a predicted portfolio riskiness :

$$branchrisk_{ot} = \sum_{i \in \text{Industries}} \frac{\text{Value of Branch lending to Industry}}{\text{Value of all Branch Lending at time t}} * sharelate_i$$

Note that $branchrisk_{ot}$ is not determined by the share of loans late at branch o: rather, it measures the share of loans with late repayment predicted by the industrial composition of lending.

Table 8 reports how $branchrisk_{ot}$ varies following the discovery of an alleged fraud. Column (2) presents equation 5 with branch and year fixed effects. Two years prior to the discovery of fraud, the measure of risk exposure is no different than branches in which a fraud is not about to be discovered. One year prior to the fraud, the level of risk taking is elevated. However, one year following the discovery of fraud, the riskiness of the loan portfolio drops substantially, with $branchrisk_{ot}$ falling approximately eight percentage points. The standard deviation of branchrisk$_{ot}$ is approximately seven percentage points, so this represents a significant drop.

This finding is consistent with the model, which predicts that lending to old, successful firms will be increasing in the posterior probability that the firm is good type.

6.2.2 Loan Repayment

Credit drops substantially following discovery of a fraud, as does the risk-taking behavior of bank branches. What happens to the quality of loans made by an affected branch?

Because we do not observe a panel of loans, we cannot differentiate loans made prior to discovery of fraud from loans made after discovery. Instead, we focus on the growth in total credit, and the growth in bad credit.

The effect on aggregate credit is presented in columns (1) and (2) of Table 9. Columns (2) and (4) include branch fixed effects. Column (3) shows that the total amount of credit marked as late is much higher in the year in which the fraud is discovered, drops in the following year, and again increases slightly above the steady state in the following year. The increase in the year of the discovery is probably mechanical: as the CVC team investigate the bank, they identify a number of loans that are to be written off. The reduction of default in year $t+1$, followed by an increase the following year, is consistent with the ever-greening predicted by the model: in year $t+1$, the loan officer ever-green to avoid default, and this tends to reduced default (even though that coefficient is not significant). However, some of the firms that were artificially supported then fail by the following year, which led to a sharp increase in default by year $t+2$, before things go back to normal in year $t+3$. It is interesting that this time series pattern fits the model so closely.

6.3 An Attempt At Reform

On January 1, 1999, the Central Vigilance Commission, in response to criticism mentioned above, introduced a special chapter of the Vigilance Manual. The goal of this chapter was to clarify the standard and procedures of investigation of employees of public sector banks. This suggests a natural experiment: if the manual changed the incentives faced by loan officers, the effects of investigation should change.

To test this, we estimate equation 5 using two sets of D variables: one set for the period before January 1, 1999, and one set for the period after 1999. Results are presented in Table 10. The results suggest that the new vigilance chapter did not change the effect of a vigilance investigation: the β coefficients are very similar in the pre-and post-reform period. Panel B gives the cumulative effect of three years following the discovery of the fraud (the value $\beta_0 + \beta_1 + \beta_2 + \beta_{\geq 3}$). The cumulative effect prior to the reform is -8.15, which is indistinguishable from the cumulative effect following the reform, of -7.64.

The inefficacy of the manual was predicted in early 2002 by a Public Sector Bank Manager in the newspaper “The Hindu:”

> A deeper examination would clearly show that a separate manual is unlikely to improve the matter of oppressive levels of accountability in PSBs. To recall a similar Governmental supervision, in 1992 when the country was almost on the verge of insolvency and industrial growth was at low levels, one of the constraints to growth was identified as controls on creation of new industrial capacity. The then Finance Minister did not substitute one licensing policy by another, but abolished licensing requirements for all but a handful of industrial sectors. The results of this bold action are there for all to see.[8]

7 Conclusion

This paper derives and tests a model of bank lending behavior, in the face of outside supervision. We then test this model using an unusual data set from India, which combines data on almost all the credit frauds detected over two decades and the universe of loans over a long period. The data largely support the main prediction of the model: lending declines following vigilance activity, bankers shift to less risky industries, and defaults follow very specific pattern: after the initial increase due to the vigilance activity, they decline to levels below the “normal” level of default, only to increase again to high level in the following year.

These findings are confirmed using two independent datasets on vigilance activity, and in monthly, quarterly, and annual fraud data. The most precise estimate comes from quarterly branch loan data: immediately following an investigation, lending falls 4% relative to non-affected branches, and this effect is quite persistent, lasting over two years. The effect occurs primarily at the affected branch, but spills over to other branches of the same bank in the same credit market (which appears to be also affected by heightened vigilance. Other bank branches in the market did not increase lending to “replace” the missing credit. There is no systematic increase (or decrease) in credit prior to discovery of a fraud, implying that the observed effects may have a causal interpretation. Thus, in a very real sense, the bank officers are correct: fear of prosecution causes lending officers to act substantially more cautiously than their peers, who

[8] R. Viswanathan, The Hindu, Jan 12, 2002

are not affected by recent discovery of fraud.

Determining whether these effects are efficient or inefficient is more difficult. We present suggestive evidence that vigilance activity is linked "under-lending." Given that prior research provides evidence of credit constraints in precisely this context, vigilance activity may indeed be inefficient. First, the fact that vigilance activity affects aggregate lending at the bank level is hard to reconcile a bank-wide slowdown in credit following the charge of a single (albeit prominent) individual. Second, the effect is found in other branches belonging to the same bank in the same town. Third, the size of the credit declines are much larger than the amount of money involved in the fraud, and are very persistent. Finally, we see that the lending strategy taken by the affected branch changes substantially. Rather than "root out" a corrupt officer and continue to lend to the optimal mix of borrowers, bank branches affected by a vigilance activity shift lending towards safer industries.

8 References

Banerjee, Abhijit, Shawn Cole, and Esther Duflo, 2005. “Bank Financing in India,” in India’s and China’s Recent Experience with Reform and Growth, Wanda Tseng and David Cowen, eds., Palgrave MacMillion: New York, New York.

Banerjee, Abhijit, and Esther Duflo, 2004. “Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program,” working paper, MIT.

Banerjee, Abhijit, and Esther Duflo, 2006. “Growth Theory Through the Lens of Development Economics,” Elsevier Handbook of Macroeconomics.

Beck, Thorsten, Asli Demirguc-Kunt, and Ross Levine, “Bank Supervision and Corruption in Lending,” forthcoming, *Journal of Monetary Economics.*

Central Vigilance Commission 2000, “Cases Referred to the Central Vigilance Commission by Public Sector Banks: A Critical Analysis,” Central Vigilance Commission, Delhi.

Central Vigilance Commission 2001, “Special Chapter on Vigilance Management in Public Sector Banks vis-à-vis The Role and Functions of the CVC” Central Vigilance Commission, Delhi.

Cole (2008) “Fixing Market Failures or Fixing Elections: Agricultural Credit in India,” forthcoming, American Economic Journals: Applied Economics.

Dinc, Serdar (2006). “Monitoring the Monitors: The Corporate Governance in Japanese Banks and Their Real Estate Lending in the 1980s,” Journal of Business, 79(6): 3057-3081.

Galor, O. and J. Zeira (1993). ”Income Distribution and Macroeconomics.” *Review of Economic Studies* 60(1): 35-52.

Gupta, Manash, and Sarbajit Chaudhuri, “Formal Credit, Corruption and the Informal Credit Market in Agriculture: a Theoretical Analysis,” *Economica* (64):331-43.

Khwaja, Asim and Atif Mian (2005). “Do Lenders Favor Politically Connected Firms? Rent Provision in an Emerging Financial Market” *The Quarterly Journal of Economics,* 120(4): 1371-1411.

Khwaja, Asim, and Atif Mian (2007). “Tracing the Impact of Bank Liquidity Shocks: Evidence from an Emerging Market,” forthcoming, American Economic Review.

de Mel, Suresh, David McKenzie, and Christopher Woodruff, 2006. “Returns to Capital in Microenterprises: Evidence from a Field Experiment," Working paper, World Bank.

Hertzberg, Andrew, Jose Liberti, and Daniel Paravasini (2007). "Information and Incentives Inside the Firm: Evidence from Loan Officer Rotation," manuscript, Columbia University.

Mookherjee, Dilip (2006), "Decentralization, Hierarchies, and Incentives: A Mechanism Design Perspective."

Rajan, Raghu (2004). "Why Bank Credit Policies Fluctuate: A Theory and Some Evidence," Quarterly Journal of Economics, Vol. 109, No. 2: 399-441.

Tannan, M, 2001, "Banking Law and Practice in India," India Law House: New Delhi.

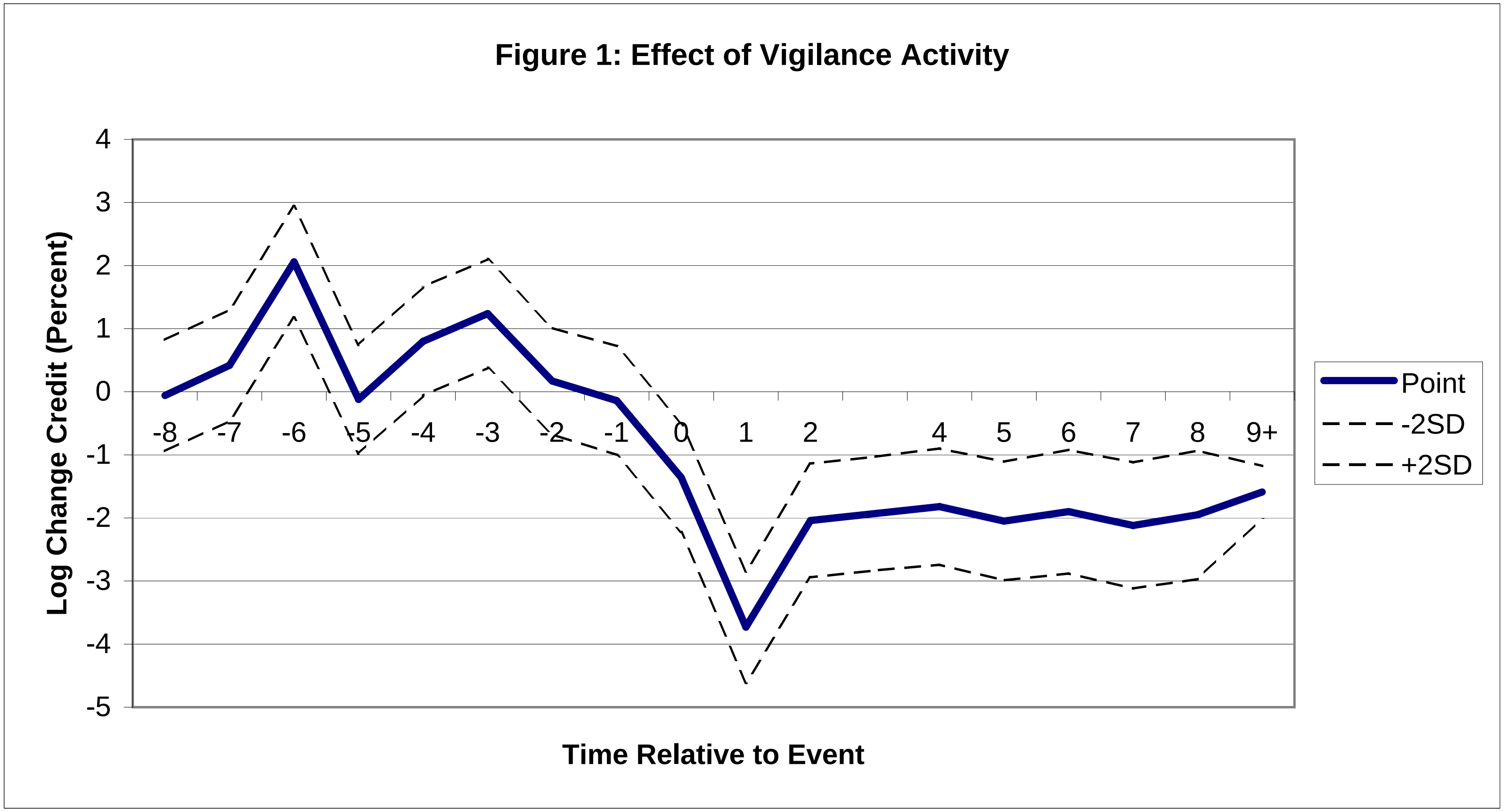

Figure 1: Effect of Vigilance Activity
Log Change Credit (Percent)
4
3
2
1
0
-1
-2
-3
-4
-5
-8
-7
-6
-5
-4
-3
-2
-1
0
1
2
4
5
6
7
8
9+
Time Relative to Event
Point
-2SD
+2SD

Table 1: Summary Statistics

Types of Fraud	**Total**
Extension of credit for illegal gratificaton	323
Other credit-related fraud	539
Total	862
Share of Frauds in Government Banks	
Extension of credit for illegal grantification	0.9
Credit-Related Other	0.8
Size of Fraud (Occurrence)	**Overall**
Average Size (Rs. 1000s)	506.4
Average Size (US Dollars 1000s)	14.5
Minimum (USD)	0
25th Percentile (USD)	94
Median (USD)	312
75th Percentile (USD)	1,314

This table contains summary statistics for the Reserve Bank of India fraud da

Table 1, Continued

		N	Mean	Std. Dev
Panel B:	**Credit Summary Statistics**			
Quarterly	Jan '80 - March 2006			
	Change in Log Credit	2,493,537	3.50%	2.78%
Annual	1992 - 2005			
	Change in Log Credit	405,227	12.45%	
	Change in Log (Late Credit)	245,774	-3.04%	
	Share of Loans Late (Value-Weighted)	365,909	15%	
	Distribution of Loan Sizes (USD 1000s)	**1992**		
	Average	$598		
	Minimum	$0		
	25th Percentile	$61		
	Median	$105		
	75th Percentile	$170		

Table 2: Where are Frauds Detected?

Panel B: Conditional Correlations

	Fraud Detected at Branch		
	(1)	(2)	(3)
Log Branch Size	0.60 ***	0.63 ***	0.75 ***
	(0.06)	(0.06)	(0.08)
Share of Credit Nonperforming	0.48 *	0.39	0.77
	(0.27)	(0.29)	(0.51)
Population Density in Credit Market	0.33 ***	0.30 ***	
	(0.07)	(0.08)	
Number of Bank Branches	0.00 ***		
	(0.00)		
Log Size of Credit Market	-0.16 **	-0.22 ***	
	(0.07)	(0.08)	
State Corruption Index	-24.63 **		
	(11.98)		
Fixed Effects	None	District	Credit Market

Notes: Each column represents a regression. The unit of observation is a bank branch (there are 44,547 branches (all government-owned banks). The dependent variable is equal to 100 if a fraud was discovered at that branch between 1980 and 2005. The dataset identifies 862 frauds.

Table 3: Serial Correlation in Investigation

	Dependent Variable: Fraud Detected in Branch This Year						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Fraud detected in branch last year	0.023 ***	0.024 ***	0.025 ***	0.026 ***	0.027 ***	0.017 **	0.017 **
	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)
Fraud detected in branch 2 years ago		0.012 **	0.012 **	0.012 **	0.011 **	0.010 *	0.010 *
		(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Fraud detected in branch 3 years ago			0.019 ***	0.020 ***	0.019 **	0.018 **	0.018 **
			(0.007)	(0.008)	(0.008)	(0.008)	(0.008)
Fraud detected in branch 4 years ago				-0.002	-0.002	-0.003	-0.003
				(0.003)	(0.003)	(0.003)	(0.003)
Fraud detected in branch 5 years ago					-0.001	-0.002	-0.002
					(0.003)	(0.003)	(0.003)
Fraud detected *bank-town* in previous year						0.011 ***	0.009 ***
						(0.001)	(0.001)
Fraud detected in *town* in previous year							0.002 ***
							(0.000)
N	641,228	595,426	549,624	503,822	549,624	447590	447590

Panel A is a regression consisting of 45,000 branches from 1990 to 2005. The dependent variable, y_{bt}, is set to 1 if a fraud was detected in branch b in year t, and 0 otherwise.

The dependent variables in columns (1) - (5) are the lagged dependent variable, indicating whether a fraud was detected in a particular branch n years prior.
Panel A is a regression consisting of 45,000 branches from 1990 to 2005. The dependent variable, ybt, is set to 1 if a fraud was detected in branch b in year t, and 0 otherwise.
Column (7) includes in addition a dummy variable indicating whether any other bank branch in the same town was investigated in the previous year.
All regressions include log credit in 1992 of the branch, and log credit squared, as controls.

Table 4: Effect of Discovery of Fraud on Lending
All Frauds, Regressions at Branch Office Level

Time Window	**Log Change**	**Log Change**	**Log Change**	**Log Change**
	(1)	**(2)**	**(3)**	**(4)**
Quarter -4	1.72 ***	1.38 ***	1.11 *	0.8 *
	(0.56)	(0.43)	(0.57)	(0.44)
Quarter - 3	1.61 ***	1.83 ***	0.96	1.24 ***
	(0.55)	(0.43)	(0.57)	(0.44)
Quarter - 2	1.69 ***	0.74 *	1.05 *	0.17
	(0.55)	(0.43)	(0.56)	(0.43)
Quarter - 1	0.21	0.41	-0.41	-0.14
	(0.55)	(0.43)	(0.57)	(0.44)
Fraud Detected	0.04	-0.75 *	-0.65	-1.36 ***
	(0.56)	(0.44)	(0.58)	(0.44)
Quarter + 1	-3.11 ***	-3.13 ***	-3.8 ***	-3.73 ***
	(0.57)	(0.44)	(0.58)	(0.45)
Quarter + 2	-0.47	-1.48 ***	-1.11 *	-2.04 ***
	(0.58)	(0.45)	(0.59)	(0.46)
Quarter + 3	-1.96 ***	-1.39 ***	-2.6 ***	-1.93 ***
	(0.59)	(0.45)	(0.60)	(0.46)
Quarter + 4	-0.31	-1.29 ***	-0.93	-1.82 ***
	(0.60)	(0.47)	(0.62)	(0.47)
Quarter + 5	-2.05 ***	-1.54 ***	-2.65 ***	-2.05 ***
	(0.61)	(0.48)	(0.63)	(0.48)
Quarter + 6	0.55	-1.47 ***	0.04	-1.9 ***
	(0.63)	(0.49)	(0.65)	(0.50)
Quarter + 7	-2.66 ***	-1.73 ***	-3.14 ***	-2.12 ***
	(0.64)	(0.50)	(0.66)	(0.51)
Quarter + 8	-1.49 **	-1.55 ***	-1.97 ***	-1.95 ***
	(0.66)	(0.51)	(0.68)	(0.52)
Quarter > 8	-1.17 ***	-1.36 ***	-1.36 ***	-1.59 ***
	(0.13)	(0.10)	(0.27)	(0.21)
N	2493542	2493542	2493542	2493542
Effect Identified for:	Branch	Branch	Branch	Branch
Fixed Effects	None	Quarter	Branch	Quarter & Branch
Time Period	1990-2005	1990-2005	1990-2005	1990-2005

Notes:

Each column represents a separate regression. The dependent variable, a measure of credit at the branch level, is indicated above the regression.

Column (1)-(5) use change in log credit as the dependent variable.

The regressions measure the effect of discovery of fraud on lending at the branch level.

The approach is a difference-in-difference or "event study" methodology.

"Fraud Detected" is a dummy variable taking the value of 1 if a fraud was detected at a specific branch in a specific quarter.

The "Quarter + N" dummy is equal to 1 at time t if a fraud was detected at that branch at time t-N.

The "Quarter - N" dummy is equal to 1 at time t if a fraud was detected at that branch at time t+N.

Columns 2 includes quarter fixed effects; Column 3 includes branch level fixed effects, and Column 4 includes branch and quarter fixed effects.

Table 5: Longer-Term Effect of Fraud Discovery

Time Window	Branch Level Effect
	(1)
Quarter - 8	-0.533
	(0.487)
Quarter - 7	-0.337
	(0.489)
Quarter - 6	1.841 ***
	(0.500)
Quarter - 5	0.033
	(0.503)
Quarter -4	0.727
	(0.517)
Quarter -3	1.022 *
	(0.530)
Quarter - 2	0.300
	(0.543)
Quarter - 1	0.444
	(0.555)
Fraud Detected	0.226
	(0.574)
Quarter + 1	-3.074 ***
	(0.586)
Quarter + 2	-2.102 ***
	(0.613)
Quarter + 3	-1.655 ***
	(0.630)
Quarter + 4	-1.412 **
	(0.650)
Quarter + 5	-1.470 **
	(0.670)
Quarter + 6	-1.750 ***
	(0.680)
Quarter + 7	-1.870 ***
	(0.690)
Quarter + 8	-1.710 **
	(0.700)
Quarter + 9	-2.170 ***
	(0.720)
Quarter + 10	-1.660 **
	(0.750)
Quarter + 11	-2.380 ***
	(0.760)
Quarter + 12	-0.790
	(0.770)
Quarter + 13	0.390
	(0.780)
Quarter + 14	-1.730 **
	(0.800)
Quarter + 15	-1.490 *
	(0.810)
Quarter + 16	-2.070 ***
	(0.800)
Quarter > 16	-0.920 **
	(0.460)
Fixed Effects	Quarter & Branch
N	1421749

Notes:

The regressions measure the effect of discovery of fraud on lending at the branch level.

The approach is a difference-in-difference or "event study" methodology.

The "Quarter + N" dummy is equal to 1 at time t if a fraud was detected at that branch at time t-N.

The "Quarter - N" dummy is equal to 1 at time t if a fraud was detected at that branch at time t+N.

Table 6: Spillover Effects of Fraud on Lending

Time Window	Affected branch	Non-affected branch of bank	Non-affected branch in same town
	(1)	(2)	(3)
	0.774	-0.085	-0.103
	(0.535)	(0.137)	(0.063)
Quarter - 3	0.979 *	-0.161	0.148 **
	(0.547)	(0.140)	(0.063)
Quarter - 2	-0.291	0.141	0.385 ***
	(0.560)	(0.142)	(0.063)
Quarter - 1	0.097	0.147	0.237 ***
	(0.571)	(0.143)	(0.063)
Fraud Detected	-0.060	0.267 *	0.007
	(0.592)	(0.148)	(0.064)
Quarter + 1	-2.621 ***	-0.712 ***	0.180 ***
	(0.605)	(0.151)	(0.063)
Quarter + 2	-2.380 ***	0.136	0.110 *
	(0.631)	(0.157)	(0.063)
Quarter + 3	-1.642 **	0.011	-0.197 ***
	(0.650)	(0.159)	(0.064)
Quarter + 4	-1.585 **	0.056	-0.080
	(0.670)	(0.163)	(0.065)
Quarter + 5	-1.740 **	-0.021	0.240 ***
	(0.690)	(0.168)	(0.064)
Quarter + 6	-1.720 **	-0.264	0.141 **
	(0.710)	(0.174)	(0.064)
Quarter + 7	-2.330 ***	0.326 *	0.034
	(0.710)	(0.179)	(0.064)
Quarter + 8	-2.010 ***	0.149	0.077
	(0.730)	(0.186)	(0.064)
Quarter + 9	-1.950 ***	-0.480 **	0.135 **
	(0.750)	(0.192)	(0.064)
Quarter + 10	-2.020 ***	0.418 **	-0.226 ***
	(0.770)	(0.197)	(0.064)
Quarter + 11	-2.860 ***	0.794 ***	-0.548 ***
	(0.790)	(0.197)	(0.066)
Quarter + 12	-0.900	-0.056	-0.014
	(0.790)	(0.194)	(0.065)
Quarter + 13	0.150	-0.256	0.415 ***
	(0.800)	(0.193)	(0.065)
Quarter + 14	-2.020 **	0.226	-0.134 **
	(0.830)	(0.200)	(0.066)
Quarter + 15	-2.150 **	0.382 *	0.293 ***
	(0.840)	(0.205)	(0.066)
Quarter + 16	-2.360 ***	-0.087	0.118 *
	(0.870)	(0.208)	(0.067)
Quarter > 16	-0.970 **	-0.080	-0.311 ***
	(0.480)	(0.130)	(0.071)
Fixed Effects	Quarter & Branch		
R^2	0.080		
N	1421538		

Note: All three columns present coefficients from **one** regression. The coefficients are estimates from the event window, as defined by the ϕ function in the text.

The dependent variable in column 1 is changes in log credit from affected branches.

The dependent variable in column 2 is changes in log credit from non-affected branches of same bank in same town.

The dependent variable in column 3 is changes in log credit from non-affected branches of non-affected bank in town.

Table 7: Credit Using Annual Data

Time Window	(2)	
Year - 2	1.1	
	(1.03)	
Year - 1	-0.22	*
	(1.11)	
Fraud Detected	-2.27	***
	(1.21)	
Year + 1	-5.43	***
	(1.38)	
	-6	***
	(1.58)	
Year >2	-6.67	***
	(1.55)	
R2	0.26	
N	262401	
Fixed Effects	Year & Branch	

Notes:
Each column represents a separate regression.
The dependent variable is change in log credit at the branch level.

Table 8: Risk Appetite and Discovery of Fraud

Time Window	**All Credit Definition**	
	(1)	(2)
Year - 2	-1.16	-1.3
	(2.44)	(2.42)
Year - 1	7.15 **	7.28 **
	(3.46)	(3.43)
Fraud Detected	-2.14	-2.12
	(2.28)	(2.26)
Year + 1	-7.76 **	-8.04 **
	(3.94)	(3.91)
	4.61	4.54
	(3.08)	(3.06)
Year >2	-1.27 ***	-0.88 ***
	(0.04)	(0.12)
R2	0.47	0.48
N	408555	398990
Fixed Effects	Branch	Year & Branch

Table 9: Bad Credit and Discovery of Fraud

Time Window	Change in Log Credit		Change in Log Bad Credit	
	(1)	(2)	(3)	(4)
Year - 2	0.85	0.79	-0.64	-0.45
	(0.82)	(0.92)	(1.36)	(1.65)
Year - 1	0.08	-0.42	-0.54	-0.98
	(0.85)	(0.97)	(1.39)	(1.72)
Fraud Detected	-2.42 ***	-3.03 ***	4.39 ***	5.99 ***
	(0.89)	(1.03)	(1.53)	(1.88)
Year + 1	-4.03 ***	-4.66 ***	-2.01	-1.58
	(0.97)	(1.15)	(1.59)	(2.05)
	-4.52 ***	-5.33 ***	4.37 ***	4.09 *
	(1.03)	(1.24)	(1.66)	(2.18)
Year >2	-3.48 ***	-4.72 ***	-0.81	-0.75
	(0.39)	(1.09)	(0.57)	(1.88)
R2	0.04	0.2	0.01	0.29
N	357266	357266	154139	154139
Fixed Effects	Year	Year & Branch	Year	Year & Branch

Notes:

Each column represents a separate regression.
The dependent variable is change in log credit at the branch level, change in log bad credit or share of bad loans.
The regressions measure the effect of discovery of fraud on lending at the branch level.
The approach is a difference-in-difference or "event study" methodology.

"Fraud Detected" is a dummy variable taking the value of 1 if a fraud was detected at a specific branch in a specific year.
The "Year + N" dummy is equal to 1 at time t if a fraud was detected at that branch at time t-N.
The "Year - N" dummy is equal to 1 at time t if a fraud was detected at that branch at time t+N.

Table 10: Effect Prior to and Following the Introduction of Special Vigilance Manual

Panel A: Regression Results		All		Public	
		(1)		**(2)**	
PRE-REFORM	Year -2	0.15		0.16	
		0.4		0.41	
	Year -1	0.1		0.47	
		0.44		0.46	
	Fraud Detected	-1.06		-0.68	
		0.85		0.89	
	Year + 1	-3.03	***	-2.95	***
		0.49		0.51	
	Year + 2	-1.86	***	-2.15	***
		0.55		0.57	
	Year >=3	-2.2	***	-2.05	***
		0.32		0.33	
POST-REFORM	Year -2	0.67	**	0.8	***
		0.29		0.3	
	Year -1	0.61	**	0.48	
		0.27		0.28	
	Fraud Detected	-1.51	***	-0.88	
		0.52		0.54	
	Year + 1	-2.28	***	-2.34	***
		0.28		0.29	
	Year + 2	-2.18	***	-2.24	***
		0.3		0.31	
	Year >=3	-1.67	***	-1.55	***
		0.23		0.24	
	R^2	0.42		0.43	
		.		.	
	N	2493609		2260733	

Panel B: Cumulative Effect			
PRE	Cumulative Effect	-8.15	-7.83
POST	Cumulative Effect	-7.64	-7.01

APPENDIX: Main Results with Clustering

	Unadjusted Errors (1)	Clustered by Branch (2)
Quarter - 8	-0.060	0.195
	(0.450)	(0.584)
Quarter - 7	0.420	0.425
	(0.450)	(0.571)
Quarter - 6	2.060 ***	2.412 ***
	(0.450)	(0.570)
Quarter - 5	-0.120	1.124 *
	(0.440)	(0.608)
Quarter - 4	0.800 *	1.116 *
	(0.440)	(0.675)
Quarter - 3	1.240 ***	0.964
	(0.440)	(0.662)
Quarter - 2	0.170	1.052
	(0.430)	(0.661)
Quarter - 1	-0.140	-0.401
	(0.440)	(0.675)
Fraud Detected	-1.360 ***	-0.639
	(0.440)	(0.694)
Quarter + 1	-3.730 ***	-3.785 ***
	(0.450)	(0.684)
Quarter + 2	-2.040 ***	-1.093
	(0.460)	(0.698)
Quarter + 3	-1.930 ***	-2.583 ***
	(0.460)	(0.737)
Quarter + 4	-1.820 ***	-0.913
	(0.470)	(0.746)
Quarter + 5	-2.050 ***	-2.641 ***
	(0.480)	(0.760)
Quarter + 6	-1.900 ***	0.049
	(0.500)	(0.829)
Quarter + 7	-2.120 ***	-3.123 ***
	(0.510)	(0.850)
Quarter + 8	-1.950 ***	-1.956 **
	(0.520)	(0.852)
Quarter > 8	-1.590 ***	-1.338 ***
	(0.210)	(0.252)
Fixed Effects	Quarter & Branch	Quarter & Branch
Clusters		55390
N	2493542	2493542

Notes:

"Fraud Detected" is a dummy variable taking the value of 1 if a fraud was detected at a specific branch in a specific quarter.

The approach is a difference-in-difference or "event study" methodology.

"Fraud Detected" is a dummy variable taking the value of 1 if a fraud was detected at a specific branch in a specific quarter.

The "Quarter + N" dummy is equal to 1 at time t if a fraud was detected at that branch at time t-N.

The "Quarter - N" dummy is equal to 1 at time t if a fraud was detected at that branch at time t+N.

In Column 2, standard errors are adjusted for correlation over time in each particular bank branch, using the SAS GENMOD command