RATIONAL HERDING AND THE SPATIAL CLUSTERING OF BANK BRANCHES: AN EMPIRICAL ANALYSIS

Angela Chang, Shubham Chaudhuri, and Jith Jayaratne

Federal Reserve Bank of New York
Research Paper No. 9724

August 1997

Rational herding and the spatial clustering of bank branches: an empirical analysis*

Angela Chang

Shubham Chaudhuri
Department of Economics
and
School of International and Public Affairs,
Columbia University

Jith Jayaratne
Federal Reserve Bank of New York

August 1997

Abstract

Bank branches in New York City tend to be spatially clustered. For instance, of the 221 branches that were opened in New York City between July, 1990 and June, 1995, 181 (or 82 percent) were opened in census tracts that already had at least one other branch. A number of recent theoretical papers have highlighted the possibility of rational herding in various arenas of economic activity. This paper explores empirically whether the apparent clustering of bank branches can be at least partially attributed to rational herding by banks. We find that even after controlling for the expected profitability of operating a branch in an area, branch openings follow other, existing branches. Moreover, such bandwagon behavior appears to reduce branch profits. These findings, combined, suggest that herd behavior may be a factor in the branch location decisions of banks.

**PRELIMINARY. PLEASE DO NOT QUOTE. Comments welcome. The views expressed in this paper are those of the authors and do not necessarily reflect the opinions of the Federal Reserve Bank of New York or of the Federal Reserve System. Corresponding author: Shubham Chaudhuri, Department of Economics, Columbia University MC-3308, New York, NY 10027; e-mail: sc301@columbia.edu.

1. Introduction

Bank branches in New York City (and in other metropolitan areas) tend to be spatially clustered. For instance, of the 913 bank branches that were in operation in New York City in June, 1990, 66 percent were located in census tracts where there was at least one other branch even though 79 percent of the census tracts had no branches. Moreover, of the 221 branches that were opened in New York City between July, 1990 and June, 1995, 181 (or 82 percent) were opened in tracts that already had at least one other branch. The aim of this paper is to explore empirically whether the apparent clustering of bank branches can be at least partially attributed to rational herding by banks.

The term "rational herding" has been used to describe situations in which it is individually rational for agents/firms to mimic the actions of others even though such mimicry can potentially lead to aggregate outcomes that are sub-optimal.[1] A number of recent theoretical papers have highlighted the possibility of rational herding in various arenas of economic activity. These models have been used to explain stylized facts about the clustering of retail stores, patterns of technology adoption, voter choice and even fertility decisions. The idea that imitative behavior can be both individually rational and socially inefficient has intrinsic intuitive appeal. But there have been relatively few attempts to empirically test these models and formal statistical evidence of rational herding is rare. In this paper, we attempt such a test.

We focus on bank branch location for two reasons. First, branching offers several advantages as an arena in which to test for rational herding. As indicated above, bank branches tend to be spatially clustered. Moreover, the branch location decision appears to have many of the ingredients that theoretical models suggest are conducive to rational herd behavior, suggesting that such behavior may drive branch location decisions. To begin with, there is considerable uncertainty about the profitability of opening a branch in any given neighborhood and uncertainty about the right course of action is a prerequisite for most types of rational herding. Next, the costs of setting up a branch are substantial, as are the costs, both direct and indirect of closing a branch. These costs suggest that banks are at least partially locked in to the locations they choose for their branches and this makes it more likely that herding, if it exists, can be detected. Third, the branch location choice represents a discrete action—to enter or not enter a neighborhood—and the discreteness of the action space has been emphasized in some herding models. And lastly, the fact that banks generally expand their networks of branches at different times means that in deciding where to locate their branches, banks have an opportunity to observe where other banks have located their branches. These features of the branch location decision do not, of course, imply that rational herding will occur; but they suggest that it might.

If herding occurs in the location of bank branches, branch data are particularly useful in detecting such behavior. Any test of herding must separate those cases

[1] Note that this definition of herding is quite specific in that it excludes situations in which agents act *independently but similarly* as well as situations in which imitative behavior is *both individually and socially efficient*. We use the more neutral term "clustering" to refer, in a purely descriptive sense, to situations where agents appear to be taking similar actions but may or may not be herding.

where agents behave similarly because they receive identical public information from the instances where agents mimic others who preceded them. This requires a great deal of knowledge about the information available to agents. For instance, testing for herding in financial markets may be difficult because the investigator may not know much about what information was commonly available to all the agents involved. But banks rely primarily on the limited information in the Census and other public data sources when locating branches, and this information is also available to us. Even if all banks have access to some information that we lack, and we allow for this possibility in our empirical work, our point here is merely that relative to other industries, we can better control for the information available to banks in the context of branching.

A second reason we focus on bank branch location is that understanding the factors underlying the branch location decisions of banks is itself of policy interest. Bank branches tend to be unevenly distributed. This unevenness has attracted considerable attention, both in the popular media and in policy circles because community groups argue that a bricks-and-mortar branch presence is important for access to banking services. They point out, for example, that retail customers who need to cash checks and open savings accounts have few good substitutes for banks.[2]

As a result, banks face pressure to maintain a branch presence in "underserved areas." For example, in several recent bank mergers, the acquirer promised to not close existing branches in low-income neighborhoods. And, in a landmark decision in 1994, the Department of Justice announced a consent decree with Chevy Chase Federal Savings Bank in Maryland (which had most of its branches in relatively affluent neighborhoods of Washington, D.C.), whereby that bank agreed to open several offices in minority neighborhoods of Washington, D.C. (Banking Policy Report (1994)).

Such interventions, whether in the form of public pressure exerted indirectly through the regulatory process governing mergers, or direct ones of the type faced by Chevy Chase Bank may be warranted if the existing spatial distribution of bank branches reflects an underlying market failure and is therefore, in some sense, socially sub-optimal.[3] There is, however, little agreement on whether that is the case. The uneven distribution of branches partly reflects the uneven distribution of profitable opportunities. Branches may be clustered simply because the underlying demand for banking services is clustered.

However, the possibility that banks may discriminate against certain neighborhoods and individuals and refuse to provide banking services to such individuals and neighborhoods ("redlining") has received considerable attention recently.[4] Such behavior, if it exists, could produce a distribution of bank branches that is more skewed

[2]For example, check-cashing outlets often charge fees of up to $9 to cash a $500 payroll check (Caskey (1991)). Glassman (1995) provides an opposing point of view, arguing that there are a number of alternative service providers and that "banks are not necessarily the only—or the best—source of financial services for low-income communities."

[3]The appropriate form of intervention would, of course, still depend on the nature of the underlying market failure. It is also important here to distinguish between a situation where the distribution is sub-optimal from a purely efficiency perspective—the source of the inefficiency being a market failure in the banking sector—and one where the criterion for optimality is somewhat broader and includes equity considerations. The two have at times been confused in the policy debate.

[4]Tootell (1996) is the most recent example, and contains references to previous research.

than the distribution of demographic and economic factors that affect branch profitability, and may justify policy intervention.

In this paper, we illustrate the possibility that a *different* type of market failure, rational herding based on, for example, information externalities, may provide a partial explanation for the uneven distribution of branches. This type of market failure implies a different policy intervention to promote a more even distribution of branches, namely a subsidy for opening branch offices in thinly branched areas. There has been surprisingly little empirical work on the factors underlying the branch location decisions of banks. Nearly all studies of redlining have focused on lending. Avery (1991) is one of the few papers we have been able to identify that directly examines the branch location choices of banks. Nor has there been, to the best of our knowledge, any discussion of alternative explanations such as rational herding for the uneven distribution of bank branches.[5] This paper aims to fill both these gaps.

We propose the following simple test of herd behavior. Controlling for the expected profitability of operating in a given tract, the probability of a branch being opened in a tract should not, in the absence of herding, increase with the number of existing branches in that tract (or should depend negatively to the extent that competition is tougher in neighborhoods with a large number of branches). We test this hypothesis using a new, extensive dataset on New York City census tracts for the 1990-95 period. We find that, consistent with the hypothesis of herding, banks are more likely to open branches in tracts where there are already other branches, *ceteris paribus*.

We then test the robustness of this finding and clarify its interpretation. This further test is motivated by our concern that the statistically significant positive relationship between branch openings and the number of existing branches that we take as evidence of herding may be spurious, arising instead from unobserved (to us, but not to the banks) determinants of profitability. Using deposits as a proxy for branch profitability, we find that profits decrease when banks follow other banks' branches. This suggests that the observed pattern of branch openings following existing branches cannot be explained in terms of existing branches proxying for unmeasured determinants of profitability. The fact that average deposits per branch decrease with the number of branches in a tract also suggests that the herd behavior that we document stems from either informational externalities or reputational concerns rather than from positive locational externalities, e.g., agglomeration externalities due to consumer search behavior.

The next section provides some background on branch banking and presents evidence on the spatial clustering of bank branches and branch openings—the starting point for our analysis. In the following section, Section 3, we discuss the literature on rational herding and describe how it might explain the clustering of bank branches. Section 4 outlines our empirical strategy and provides a description of the data we use. In Section 5, we present the evidence on herding. Section 6 discusses alternative interpretations of the evidence and their implications for policy. Section 7 concludes.

[5] An exception is Lang and Nakamura (1993) which presents a model (that we discuss in more detail later) of mortgage redlining that yields herding based on a dynamic information externality.

2. Preliminaries

2.1. Branch banking

Despite the growth, in the last fifteen years, of a number of alternative mechanisms for delivering banking services (such as ATMs, phone banking, PC banking, and centralized loan originations), banks continue to rely on traditional, brick-and-mortar branches.[6] The primary reason for this is that though ATMs and phone banking are widely used, their usage is typically limited to specialized functions such as information inquiries and withdrawals. Bank customers continue to use branches to make deposits. For example, a 1995 Master Card Survey of major retail banks found that nearly 90 percent of all deposits are done in branches (Mead (1997)).

Anecdotal evidence suggests that banks investigate potential branch locations carefully. They often hire market survey firms to produce site studies. Moreover, banks appear to use a fine geographic grid when scouting for branch location sites, suggesting that they do not believe that locating a branch just anywhere in a city will do. A prominent market survey firm that helps banks locate branches informs us that client banks typically define a "trade area" for a branch to consist of 2000-2500 households. This is not much larger than the typical New York City census tract, which had 1253 households in 1990. New York City census tracts cover a very small geographic area, often no more than a few square blocks.

Banks may find it important to locate branches carefully for several reasons. The first is that customers appear to value proximity to a branch. A 1996 *American Banker* survey showed that the majority of bank customers who switched banks did so because they wanted to be closer to a branch (Kutler (1996)). Not surprisingly, the same survey found that the average bank customer visited a branch of her bank at least three times a month. When customers value proximity, banks cannot locate a branch anywhere and expect customers to use ATMs, etc. to bank over a distance.

A second reason for locating branches carefully is that branch profitability is uncertain, and there are substantial fixed costs of opening and closing branches. A banking market research firm informs us that banks are often unable to explain the wide variation in the performance of their own branches, and that the research firm is often hired by banks to determine the causes of such performance differences. As for the fixed costs of operating a branch, anecdotal evidence suggests that such expenses are considerable. The typical branch costs approximately $1.5 million to set up (mostly in real estate and construction costs).[7] Fixed costs of operating a branch (wages and maintenace costs) add approximately $1.4 million annually (Radecki, et al). Since the typical branch carries $50 million in deposits, the cost of setting up a branch represents a one-time addition of 300 basis points to the cost of deposit funds, and fixed operating costs add another 280 basis points annually.

Closing a branch is a costly process. Banks are required to submit a notice of a proposed closing with regulators no later than ninety days prior to the closing date. The required notice must include a detailed statement of the reasons for the decision to close a branch, and statistical and other information supporting the reasons. Although banks do not need regulator approval to close an unprofitable

[6] Nationwide, the number of bank and thrift offices declined only slightly between 1990 and 1995 from 84,419 in 1990 to 81,875 in 1995 (Federal Deposit Insurance Corporation (1996)).

[7] Set-up costs are estimates based on various industry sources.

branch, they face considerable pressure from community groups to keep branches open. In several recent instances, banks that were party to mergers committed themselves to retain existing branches in low-income neighborhoods.

2.2. Evidence on branch clustering

The population of New York City banks encompasses a wide range of institutions, from large money center banks to many small, retail banks. Ninety one independent banks and bank holding companies operated 844 branches in New York City in June, 1995. Of these, four large institutions (Bank of New York, Citibank, Chase Manhattan Bank and Chemical Bank) owned 490 offices. In this section we provide some basic descriptive statistics documenting the spatial clustering of both bank branches as well as branch openings in New York City. The data used to generate these descriptive statistics are described in more detail in a later section.

Table 1 depicts the spatial distribution of bank branches at the census tract level for two years, 1990 (top panel) and 1995 (bottom panel). The first column of each panel provides a breakdown of census tracts, by the number of branches in the tract as of June of the relevant year, i.e., 1990 or 1995; the second column shows the distribution of branches, by the number of branches in the tract in which the branch was located. These numbers indicate that in both 1990 and 1995, bank branches tended to be located in tracts where there were already other branches, while many tracts remained without branches (and the basic pattern is repeated in all of the years 1990–1995). For instance, looking at the top panel we find that of the 913 branches in existence in New York City as of June 1990, 66 percent were located in tracts where there was at least one other branch. Meanwhile,79 percent of the 2218 census tracts had no branches. That is, all 913 branches were concentrated in only 21 percent of all census tracts. This pattern had not changed in 1995 (as seen in the lower panel of Table 1).[8]

The spatial distribution of bank branches, observed in any given year, is the outcome of branch location decisions made by banks over an extended period of time. As such, the apparent clustering of existing bank branches may simply be the remnant of clustering in the past and need not therefore suggest that clustering is an ongoing phenomenon. More direct evidence on clustering can therefore be obtained from the spatial distribution of branch *openings*, which is depicted in Table 2.

The first column of the top panel of Table 2 provides a breakdown of census tracts by the number of branch openings in the tract between July, 1990 and June, 1995; the second column shows the distribution of branch openings, by the number of branch openings in the tract in which the branch was opened. There were 221 branch openings between July, 1990 and June, 1995, and these were concentrated in 142 (i.e., 7%) of the 2218 tracts. Of the 221 branch openings, 53 percent (or 117) took place in tract where there was at least one other branch opening during the five-year period.

The bottom panel of Table 2 provides perhaps the clearest indication of branch clustering. The first column shows the breakdown of the census tracts in which there were branch openings, by the number of branches that existed in the tract

[8]These figures probably under-estimate the degree of branch clustering because tracts with branches are likely to be themselves clustered and not uniformly distributed among the branchless tracts.

at the beginning of the period, i.e., in June, 1990. Here we see that 75 percent of all census tracts that experienced a branch opening between July, 1990 and June, 1995, already had a branch. More striking still is the fact that 82 percent of the 221 branches that were opened over this five-year period, opened in tracts with at least one other branch at the beginning of the period (see the second column).

The simplest and most obvious explanation for the clustering of bank branches documented above is that the demand for banking products and services is itself spatially clustered. It is certainly true that not all neighborhoods in New York City offer the same potential customer base for banks. And if the disparities across neighborhoods in the extent of demand is sufficient to outweigh the adverse effects of increased competition, banks—like Willie Sutton—might simply be following the money and locating in those areas with significant demand for banking services.

Table 3 reports some summary statistics that suggest that there is indeed some basis for this explanation. Tracts with existing bank branches as of June, 1990, as well as tracts in which branches were opened between June, 1990 and June, 1995, appear to be more affluent along a number of observable dimensions that are plausible indicators of the demand for banking services. For instance, the tracts that had branches (in June, 1990) had, on average, larger populations, fewer poor households, a better-educated population, higher median household income, and more workers and were, on average, more commercial. The same is true of tracts in which branches were opened between July, 1990 and June, 1995.

However, Table 3 also reveals that branches were more likely to be opened in tracts that already had more existing branches. On average, there were nearly 3 existing branches (in June, 1990) in the tracts in which branches were opened between July, 1990 and June, 1995; on the other hand, the average number of branches in the tracts in which no branches were opened during the five-year period, was less than one. This stark contrast at least raises the possibility that the observed clustering may be partly due to some form of rational herding.

3. Rational herding and the clustering of bank branches

A large literature on rational herding has emerged in recent years.[9] The literature suggests several different channels through which herding can arise. At least three of the suggested channels seem to us to be ways in which rational herding might occur in the location of bank branches. We describe them below.

In information cascade models, the possibility of herding stems from an information externality (Banerjee (1992), Bikhchandani, Hirshleifer and Welch (1992), Welch (1992)). The typical setup in these models has agents choosing from a set of actions according to a predetermined sequence. Each agent receives a conditionally independent private signal about the correct action to take and is also able to observe the actions, but not the signals, of those who preceded her. Using both her private information as well as the public information embodied in the choices of others, each agent updates her priors about the profitability of alternative actions and then chooses accordingly.

If the action space is coarse relative to the signal space, agents may not be able to adequately tailor their chosen action to reflect both their private information as

[9] Devenow & Welch (1996) and Gale (1996) provide very useful overviews of the literature.

well as the public information. They may rationally choose, then, to ignore their own information and base their decision on the public information—i.e., faced with a choice between acting upon her own private signal and imitating the choices of those who acted before, it may be optimal for an agent to choose the latter. But in doing so the agent ignores the fact that her private information is lost to those who follow her since her private information is not recoverable from her publicly observable action. This is the information externality at the heart of these models. If agents have identically distributed signals, all subsequent agents face an identical situation and consequently also choose to ignore their private information. The result is an information cascade. And depending on the initial pattern of choices, the actions of agents may well converge on the wrong choice—i.e., in rational herding.

Information cascades represent an extreme form of herding in which the actions of early agents completely dominate the private signals of later agents. Lang and Nakamura (1993) present a model of mortgage redlining in which a somewhat weaker form of herding takes place. The information externality in their model stems from the fact that the actions of predecessors affect the precision of the information available to subsequent agents. In their model, the precision of appraisals—on which mortgage lenders base the size of required down payments—depends on the volume of previous home sales in a neighborhood. Appraisals are based on the prices at which previous sales were transacted because these provide noisy signals of current property values. The higher the number of previous home sales, the more precise the appraisals, and the lower the required down payments. Lower down payment requirements in turn lead to a larger number of approved mortgage loans and hence, a larger number of current sales. The positive feedback mechanism thus generated raises the possibility of herding and sub-optimal differences in mortgage lending activity across neighborhoods.

A third possible channel through which rational herding might arise is through the reputational concerns of agents when the calibre/quality of agents is unknown. In Scharfstein and Stein (1990), one of the first models of this kind, better (informed) managers receive informative signals about the right course of action, and the errors in these signals are correlated. Uninformative signals, those received by uninformed managers, are, on the other hand, uncorrelated. The compensation (future prospects) of a manager depends on his reputation—i.e., on whether he is regarded as informed or not. In this situation, each manager has an incentive to mimic the actions of managers who have acted before him, because by doing so he maximize his chances of appearing informed. If the action results in a good outcome, he benefits; even if the action, ex-post, yields a bad outcome, the fact that other managers made a similar choice shields the manager, enabling him to 'hide in the herd', in effect, to argue that the decision was, ex-ante, an informed one. On the other hand, were the manager to act upon his private signal, where such a signal suggests a course of action different from that taken by other managers, he would run the risk of appearing uninformed if the action resulted, ex-post, in a poor outcome. Other examples of this type of herding based on reputational concerns and relative performance are provided in Zweibel (1995) and in DeCoster and Strange (1993). The latter apply the Scharfstein and Stein (1990) model to the siting decisions faced by real-estate developers concerned about their reputations with banks. The herding in these models is based on two key premises: first, that there exists an agency problem in that the incentives of decision-makers are not aligned with the

outcomes of their decisions; and second, that the compensation of agents is based in some way on relative performance standards.

Any of the three types of models described above can plausibly be applied to explain the clustering, (and possibly rational herding), of bank branches. These models assume uncertainty of outcomes and (some) irreversibility of decisions. Both conditions are observed in branching, as we noted in Section 2. All three assume that agents act sequentially in an exogenously determined order and that the actions of agents are publicly observable.[10] Banks appear, in general, to expand their branch networks at different points in time based on a number of different factors that are arguably exogenous to the branch location decision itself. And in making their branch location decisions, banks are clearly able to observe the locations chosen by other banks.[11]

Cascade models require, in addition, a discrete action space (or at least that the signal space is large relative to the action space). Incorrect cascades are prevented when the action space is fine enough for the private information of firms to be recoverable from their chosen actions. Whether one views the branch location decision as a series of binary decisions about opening or not opening a branch in each of a number of neighborhoods, or whether one views it as a single decision about the best neighborhood in which to locate a branch, branch location represents a discrete choice.[12] When a bank chooses not to open a branch in a neighborhood, or even if it does, the strength of its private information about the profit potential of the neighborhood is not revealed. All that other banks are able to observe is the discrete location choice. And because they rationally infer from this choice that the bank's private signal was not strong enough to warrant a different course of action, these other banks may choose to (not) locate branches in neighborhoods where the bank chose to (not) locate its branch. In the process, banks may ignore their private information that alternative locations are equally profitable (or even more profitable)

[10] A separate strand of the literature on informational externalities relaxes this assumption and allows agents to choose when to act (Hendricks & Kovenock (1988), Caplin and Leahy (1993), Chamley and Gale (1994), Gul and Lindholm (1995)). In these endogenous timing models, all agents have an incentive to wait because the actions of early movers provide additional information that can improve the quality of the decisions made by late movers. The resulting equilibrium resembles a war of attrition in which agents try to out-wait others and this leads to sub-optimal delays in action.

There are two main differences between these models and the ones we discussed above. The first is that the clustering of agents' actions in endogenous timing models need not be inefficient; the second is that the inefficiency always takes the form of 'underinvestment' because of excessive delay. In sequential action models, on the other hand, 'overinvestment' (excessive clustering) is also a possibility.

[11] A slightly tricky point here is whether or not banks actually observe that certain locations were *rejected* by other banks. To the extent that banks, in principle, consider all neighborhoods within the city (subject to some obvious exceptions) to be potential sites for new branches, the location choices that are actually made implicitly indicate that other sites were rejected.

[12] While it is true that branch location could be thought of as a selection from a continuum of possible sites, for this to eliminate the possibility of herding it would have to be the case that banks gain from being "close" to the true optimal site—e.g., have a payoff function that is concave in the action space (see Lee (1993)). This seems unlikely, especially in New York City, where fairly affluent neighborhoods often adjoin more depressed areas and there does not appear to be any discernible monotonicity in the geographical positioning of neighborhoods according to their level of affluence. This suggests that banks face a payoff function similar to that in Banerjee (1992), which, because of its "all-or-nothing" form, effectively discretizes the action space and thus allows the possibility of herding.

and this can result in overclustering of branches in some neighborhoods while other neighborhoods remain underserved.

If banks care about the precision of the information available to them about the profit potential of alternative neighborhoods, they may also choose to locate branches in tracts with a larger number of existing branches. The presence of one or more existing bank branches in a neighborhood can provide additional sources of information—though branch-level profit figures are usually not available, data on branch-level deposits are readily obtained—and such information, when combined with any private information that the bank has acquired through, for instance, site analysis studies can reduce the uncertainty surrounding the profitability of opening a branch in a neighborhood.

Reputational concerns may also influence the branch-location decision if the evaluation (and compensation) of managers is based partly on the ex-post relative profitability of their branch-siting decisions. From the perspective of a manager responsible for making the branch location decision, it may be much more attractive to locate a branch in a neighborhood where there are already several other existing branches than to venture into a virgin neighborhood, even when the latter appears to have significant potential. By doing so, the manager avoids the possibility that he will be blamed for poor judgement.

The discussion above has been in largely heuristic terms. We have not written down a specific structural model of bank branch location. Partly this is due to the fact that we remain agnostic, at this point, about which of these specific models applies in the case of bank branch location. A more important reason is that we cannot, at this stage, empirically distinguish between the alternative channels through which herding might be occurring. Our empirical strategy is based, therefore, on what might be considered the common reduced form implication of the three approaches outlined above—namely, that the branch location decisions of banks should be *directly* influenced by the location decisions made earlier by other banks, over and above any publicly observable direct indicators of the profit potential of a neighborhood. However, for purely illustrative purposes, we present in the Appendix a model of bank branch location that yields herding based on an information externality along the lines of Lang and Nakamura (1993). We outline our empirical strategy in more detail in the next section.

4. Empirical strategy and data

We have made a *prima facie* case for the hypothesis that rational herding provides at least a partial explanation for the clustering of bank branches described in Section 2. The main competing hypothesis, which we also noted earlier, is that the clustering of bank branches is driven *entirely* by the fact that the demand for banking services is itself clustered. In this section,we outline a simple test of branch herding that allows us to distinguish, empirically, between these two competing hypotheses. We also describe the data we use to implement the test.

4.1. A test of herding

We adopt a reduced form approach in testing for herding. Our starting point is the expression for the process generating branch-level profits:

$$\pi_{jt} = X_{jt}\alpha + \delta N_{jt} + e_{jt} \tag{4.1}$$

Here π_{jt} represents the profits from operating a branch in tract (or neighborhood) j, starting in time period t.[13] X_{jt} is a vector of demographic and other factors at time t that affect the expected profitability of operating a branch in tract j and N_{jt} is the number of existing branches in tract j at the beginning of period t. These variables are assumed to be observable to us, as well as to the banks. The disturbance term, e_{jt}, captures any factors affecting branch-level profits that we assume, for the moment (see the discussion in the next subsection), neither we nor the banks observe.

In the absence of any positive locational externalities—i.e., increased profits from locating close to other branches—N_{jt} simply proxies for the degree of bank competition in the tract, *ceteris paribus*.[14] Increased competition within a neighborhood is likely to decrease branch-level profits, and so we expect the coefficient on N_{jt} to be negative, or at least non-positive, i.e., $\delta \leq 0$.

The test of herding that we carry out is based on the simple proposition that, as long as banks base their branch location decisions on the profit potential of a neighborhood, in the absence of any herding, the reduced form expression for the number of branch openings in a tract should mirror that for branch-level profits. In particular, if we adequately control for the factors, X_{jt}, that independently affect the expected profitability of operating in tract j, the number of branch openings in a tract should, in the absence of herding, depend negatively (if at all) on the number of existing branches in the tract since more branches indicate stiffer competition.

This suggests estimating the following test equation:

$$O_{jt} = X_{jt}\beta + \gamma N_{jt} + u_{jt} \tag{4.2}$$

where the dependent variable O_{jt} is the number of branch openings in tract j during period t, and the other variables are as defined above. Under both the competing hypotheses, the coefficients, β, on the vector of profit factors, X_{jt}, should qualitatively match those in equation (4.1). In the absence of herding behavior of the sort described in Section 3, the effect of N_{jt} should also match that in (4.1)—i.e., γ should be less than or equal to zero.

On the other hand, if banks herd (i.e., between two otherwise equally attractive tracts, they choose the tract with more branches), the effect of N_{jt} on subsequent branch openings is the sum of two opposing forces. The competition effect implies that we ought to observe relatively fewer branch openings in tracts with relatively more branches at the beginning of the period. But a second, opposing influence, arises from the fact that tracts with a greater number of existing branches will attract more branches if banks herd. Although the overall effect of N_{jt} is indeterminate

[13] As the basic unit of analysis in this paper is a tract rather than a bank, to save on notation, we do not explicitly incorporate the obvious heterogeneity that exists among banks.

[14] We detail in Section 6 how such positive locational externalities might arise, and discuss how they affect the interpretation of our test of herding.

under the herding hypothesis, a positive correlation between the initial number of branches and subsequent branch openings would suggest herding behavior in branching.

Bank branch location data offer several advantages in testing for herding behavior. First, because banks rely substantially on public census data when making their branch location decisions, we are better able to control for the information available to banks, and hence are able to more cleanly identify herding.[15] This contrasts with, for instance, the study by Grinblatt, Titman and Wermers (1995) of the trading patterns of 274 mutual funds between 1975 and 1984. They report small but statistically significant comovements (i.e., buying and selling the same stock at the same time) in the quarterly stock holdings of these funds. They do not, however, control for any public information flows (e.g., earnings announcements) that may have driven these comovements. The "herding" that they report may, therefore, simply reflect the fact that fund managers responded *independently but similarly* to the arrival of common, new information.[16]

A second advantage of our data is that the order in which the firms acted is clearly indicated. Relating the branch *opening* decision to the spatial distribution of *existing* branches provides a natural way of examining the influence of early-movers on the actions of later agents. Other empirical papers on herding have often had to rely instead on *a priori* identification of "leaders" and "followers". For instance, Jain and Gupta (1987), which tests for herding in international lending by U.S. banks during the 1970s, explores whether smaller U.S. banks followed money center banks when lending abroad. They find that money center banks' portfolio allocations did not Granger-cause smaller banks' allocations, and they conclude that there is no evidence of herding in international lending. But this conclusion relies on the authors' correctly identifying ex ante the leaders and followers in international lending. If different banks acted as leaders when lending to different countries, the Jain and Gupta test may not pick up herding.

Perhaps the closest in spirit to the approach we take is Calem (1995). He conducts a test of Lang and Nakamura (1993) by regressing mortgage-loan approval rates in U.S. urban counties in 1990-91 on 1989 home sales (and other controls). Lang and Nakamura (1993) predict that past home sales should have a positive effect on current mortgage loan approval rates. Calem finds just such an effect, and he concludes that the data support the Lang and Nakamura (1993) model. However, he finds this effect only in non-minority tracts, a troubling result because the Lang and Nakamura (1993) model predicts that the information externality is strongest in areas with thin home sales (and minority areas have fewer sales). Moreover, the positive correlation between past home sales and current mortgage approval rates even in non-minority areas may be the result of serially-correlated demand shocks, a possibility Calem is unable to control for using cross-sectional data.

[15] It is, of course, still possible that despite our efforts to be comprehensive, banks have access to information that we do not. We consider this possibility in the next section when we discuss the results of our basic test of herding.

[16] Partly this is a matter of definition—i.e., how broady one defines "herding". As we mentioned at the outset, we follow the theoretical literature (see Devenow and Welch (1996)) and reserve the term "herding" for situations in which the actions of agents *directly* influence the actions of other agents, and this type of imitative behavior raises the possibility of systematically sub-optimal outcomes. The distinction is also important in another respect which is that "herd behavior" as we choose to define it potentially justifies some form of government intervention.

4.2. Data

We estimated the test equation using data on commercial bank branch openings in New York City census tracts between July 1, 1990 and June 30, 1995 (and to a more limited extent, branch openings between July 1, 1980 to June 30, 1985).[17] These data were obtained from the Federal Deposit Insurance Corporation's (FDIC) *Summary of Deposits* database, an annual series which lists the street addresses of all bank branches as of June 30th of each year. We used these street addresses to map each branch to a census tract and were thus able to obtain the number of bank branches in each census tract in each of the five years.[18] If a branch address appeared for the first time in a tract in June of a given year, we recorded that as a branch opening some time in the preceding twelve months.[19]

Our choice of a census tract as the basic geographical unit of analysis—the area j in the test equation—was based largely on data considerations. But it appears to correspond fairly well with what banks themselves use. For example, as mentioned earlier, a prominent market research and consulting firm that provides banks with site analysis services to aid their branch-location decisions reports that banks typically define the "trade area" of a branch to consist of a geographic area encompassing 2000–2500 customers. The average New York City census tract had 1253 households.

As controls for the potential profitability of operating a branch in a census tract, we use several population characteristics and indicators of business activity. These variables are described in Table 3. In addition to census tract population size, median family income, poverty rate, race and education, we include the fraction of population over the age of 65 because the supply of core deposits by the elderly may be relatively interest insensitive, making them more profitable bank customers. We include the fraction of renter-occupied housing units and the median value of housing in a tract as possible correlates of the size of the home mortgage market in an area. Median housing values may be also correlated with the real estate costs of operating a branch. We include these variables because they are strongly correlated with the number of branches in each census tract.[20] Moreover, market research firms that help banks locate branches rely on similar census information when conducting site analyses.

[17] Thrifts (savings banks and savings and loans) are excluded partly because thrifts are not perfect substitutes for commercial banks. Unlike banks, thrifts primarily make mortgage loans. Thrifts are excluded partly for data reasons; we are unable to get thrifts' branch locations for the early 1980s. If banks do in fact consider thrift branch locations when locating bank branches, dropping thrifts from the data creates a measurement error in the number of existing branches, N_{jt}, which will bias its coefficient toward zero.

Off-site ATMs are not included as branches in the FDIC's *Summary of Deposits* database. Hence, only full-service branches are included in the data and in our analysis. Dropping ATMs should not affect results here because 75 percent of ATMs are located inside branches (*American Banker*, November 30, 1996).

[18] The details of the mapping procedure are available upon request.

[19] Note that this procedure yields the *gross* number of branch openings in a tract. We use gross openings rather than net openings—i.e., gross openings minus closings—because the factors underlying closings are often quite different (see Section 2) from those influencing openings. We should point out also that if a branch changed hands but remained at the same street address we did not record this event as an opening.

[20] Based on a regression of the existing number of branches (note, not openings) in a tract on these tract-level variables. Results are available upon request.

As indicators of business activity in a tract we include the number of people working in a tract (provided by the Census Bureau in a customized data file using the Bureau's 1990 *Journey to Work* data), and a dummy indicator variable for whether the tract is a net importer of workers. Typically, we would expect commercial areas to be net importers and residential areas to be net exporters. We also include the percentage of each tract's land area that is devoted to commercial, industrial, and residential purposes using data provided by the New York City Planning Department.[21]

The demographic and business activity variables listed above describe the endowment of the census tracts as of 1990. We also include as controls, *changes* in the population characteristics of each census tract between 1980 and 1990. The potential future profits from operating in a tract depend not only on the current characteristics of the tract but also on future conditions. To the extent that past changes predict future changes, the measured changes in demographic variables should partially control for expected future demographic changes in the tract. Moreover, expectations of future changes in real estate prices should be captured by the median housing value variable if such expectations are capitalized into current prices.

5. Evidence on herding

In this section we first report the results of our basic test of herding. We find that, consistent with the hypothesis of herding, banks tend to open branches in tracts where there are already other branches, *ceteris paribus*. We then carry out a test of the robustness of this finding. This further test is motivated by our concern that the statistically significant positive relationship that we take as evidence of herding may be spurious, arising instead from unobserved (to us, but not to the banks) determinants of profitability.

5.1. Basic results

Columns (1) and (2) of Table 4 display our basic results.[22] They show that after controlling for census tract characteristics, the number of bank branches in a tract at the beginning of the period is positively correlated with the number of branch openings over the subsequent year, at least in those tracts with low-to-moderate number of branches. Column (1) report Poisson estimates, and column (2) reports ordered-Logit estimates.[23] The Poisson point estimates suggest that the number of initial branches is positively correlated with subsequent openings until the initial branch count is about fourteen. Thereafter, the two are negatively correlated.

[21] Land-use data are from the New York City Planning Department's 1995 *Land Use Data Files*. This database tracks the actual uses of real estate, not what the area is zoned for.

[22] The estimated equation differs from (4.2) only in that we have included a squared term (in the number of existing branches) to allow for possible nonlinearities.

[23] Poisson estimation seems natural here since branch openings are count data. Such estimates are also relatively easy to interpret. However, the Poisson estimator assumes that the opening of a branch in a tract does not affect the probability of subsequent openings. Since this is not true under the hypothesis of herding, we also provide Ordered Logit estimates. The dependent variable—the number of branch openings—was top coded in the Ordered Logit estimation into four categories: 0 branch openings, 1 branch opening, 2 openings, and 3 or more openings.

This is consistent with herding dominating openings behavior at relatively thinly-branched tracts. When there are many branches in a tract, stiff competition may discourage further branch openings. However, only three tracts had more than fourteen branches in 1995. Hence, herding dominates in 99.5 percent of the tracts with branches. Moreover, this "herding effect" is large. The Poisson estimates suggest that the expected number of annual branch openings in a tract with two branches is 33 percent greater than an otherwise identical tract with just one branch at the beginning of the year.

Interestingly, few of the other tract-level variables had a significant impact on the number of branch openings. This may be due to the relatively small number of branch openings and the resulting low power of the tests here. Nevertheless, the estimates indicate that banks appear to find commercial tracts, tracts that attract many commuters, and heavily populated tracts to be relatively more attractive. They appear to have found residential areas and poorer areas unattractive. Minority tracts also attract fewer branch openings, but this may be due to profitability factors that are correlated with the racial composition of a tract.

Table 5 contains some robustness tests. This table summarizes results from estimating the model in Table 4 on several sub-samples. To save space, we report only the coefficients on the initial branch count. A potential problem with the results in Table 4 arises from the fact that we pool five annual cross-sections of data when estimating the model. The error term of the model for a given tract may well be correlated over time because there may be unmeasured tract features that change little if any over back-to-back years.[24] We address this p oblem indirectly by re-estimating (4.2) as a single cross-sectional regression with the dependent variable redefined as the cumulative number of branch openings over the five-year period, 1990-1995. We regressed this variable on the tract conditions in 1990. The results we obtained are shown in the second panel of Table 5. The point estimates are comparable to those in Table 4 (reproduced in the top panel of Table 5 for comparison).

We next re-estimated the model in Table 4 excluding head office locations from the sample. The process of locating a head office may be quite different from that of locating a regular branch. For example, the potential for deposit taking from and lending in the particular neighborhood where the office is located may be less important for a head office than for a regular branch. Also, head offices are more likely to be located in a major commercial center such as midtown or downtown Manhattan. Such clustering of headoffices in heavily-branched areas may skew our results in favor of finding herding effects to the extent that we do not completely control for the commercial characteristics of a tract. The middle two panels in Table 5 show our estimates of the same location model without head offices for the 1992-95 period.[25] These results are also consistent with herding.

Finally, we re-estimated the openings model using branch openings data for the 1980-85 period (bottom two panels, Table 5). We obtained the same qualitative result of herding. The point estimates of the effect of the initial number of branches is bigger than for the 1990s, but this is probably because we had fewer controls for tract profitability for the 1980-85 period. (We do not have data on the extent of

[24] Despite this problem, we pooled the annual data because the number of openings and the number of initial branches in a tract varies over the five-year period. Pooling ensures that we use this information when estimating the correlation between openings and the initial branch count.

[25] We could reliably identify head offices only for the 1992-95 period.

commercial activity in tracts in the 1980s, nor do we have measures of the changes in the demographic variables between 1970 and 1980.)

5.2. A further test

Tables 4 and 5 indicate an apparent statistically significant positive effect of the number of existing branches on subsequent branch openings. We now investigate the possibility that this finding may be spurious. Our concern here is that N_{jt} may simply be picking up the influence of tract characteristics that we have not controlled for, but banks are able to observe. Although we have tried to be fairly comprehensive in including an extensive array of tract characteristics that might independently affect the profitability of operating in a tract, banks may be privy to more information than is available to us. For instance, banks often commission targeted market surveys of potential locations in addition to relying on publicly available information from sources such as the census.

More formally, we consider the possibility that there might be unobserved (to us) serially correlated tract-level characteristics (represented by Z_{jt}) that affect branch profitability. Again, this is a concern only in that banks might observe these characteristics whereas we do not. In other words, we imagine that the "true" process generating branch-level profits is given by:

$$\pi_{jt} = X_{jt}\alpha + \delta N_{jt} + Z_{jt} + e_{jt} \tag{5.1}$$

where:

$$Z_{jt} = \rho Z_{j,t-1} + e_{jt} \tag{5.2}$$

With banks basing their decisions on profit considerations, openings are then generated according to:

$$O_{jt} = X_{jt}\beta + \gamma N_{jt} + Z_{jt} + u_{jt} \tag{5.3}$$

Clearly, since $N_{jt} = O_{j,t-1} + N_{j,t-1} - C_{j,t-1}$(where C_{jt} is closings), given that (from (5.3)) $O_{j,t-1}$ is partly determined by $Z_{j,t-1}$, N_{jt} will be correlated with $Z_{j,t-1}$. This in turn implies that N_{jt} will be correlated with Z_{jt} if, as we have assumed in (5.2) above, Z_{jt} is serially correlated. Under this scenario, estimation of equation (4.2), which does not control for Z_{jt}, will suffer from omitted variable bias. And this may yield a positive statistically significant estimate of the coefficient on the existing number of branches even where there is no herding and the true coefficient on N_{jt} in equation (5.3) is $\gamma \leq 0$.

One possible solution for this problem is to assume that the tract information observed by the bank but not by us (Z_{jt}) is time invariant, and use our panel data to estimate an equation with tract fixed effects. Unfortunately, a fixed-effects Poisson (or Ordered Logit) estimation procedure will not produce consistent estimates of the number of initial branches because this variable is the sum of previous branch openings. (That is, the number of existing branches in any given year, being a function of lagged dependent variables (i.e., past openings), is pre-determined but not exogenous.)

Instead, we pursue an alternative "fix" by testing for the severity of the omitted variable bias as follows: Suppose that, in the estimation of (4.2), the positive coefficient on N_{jt} is being generated by the fact that N_{jt} is serving as a proxy

for unobserved/unmeasured demand factors. Then, because the branch-openings equation, in the absence of herding, mirrors the equation for profits, if we were to estimate equation (4.1), in which branch profitability is the dependent variable, the coefficient on N_{jt} should suffer from a similar omitted variable bias. In other words, estimation of (4.1) should also yield a spurious positive coefficient suggesting that the profitability of branches in a tract is positively correlated with the number of branches in the tract.

If, on the other hand, the positive correlation between branch openings and existing branches does in part reflect herding, the coefficient on N_{jt} in the branch-profits equation should be negative because of the increased competition from the larger number of branches.

We cannot directly estimate (4.1) because we do not have direct measures of branch profits. However, deposits held at a branch are a reasonable proxy for branch profits. The principal function of bank branches is to gather deposits. Branches play only a limited role in lending. Credit card loans are typically originated nationwide by centralized operations of specialized credit card banks. Similarly, mortgage applications are often processed and approved at centralized mortgage lending units, not in branches (at least in large banks).

If branches play any role in lending, it is in small business lending. Even here, branches are of limited importance. A 1995 survey by the Consumer Bankers' Association showed that while a majority of banks relied on branches to supply deposit services to small business customers, branches played a much smaller role in lending. Although 69 percent of the seventy two large banks surveyed said that they relied on branches for small business deposit services, only 26 percent said that they used branches to originate loans, and only 8 percent said that their branches underwrite loans (Allen (1995)). Anecdotal evidence confirms this pattern holds in New York. Several banks in the area solicit, process, approve and maintain small business loans from "loan centers" that cover a large area.[26]

Table 6 shows the results from estimating a regression of the average deposits per branch (in 1000s of dollars) in a census tract at time t on the number of branches in that tract at time t (and all other tract features used as controls in Table 4).[27] The dependent variable in the top panel in Table 6 is the value of total deposits at the average branch in each census tract (for those tracts with at least one branch) for 1990-95 and for 1984-85. (Data for previous years of the 1980s were unavailable.) Not all deposits are equally profitable. Transactions deposits (checking accounts) are typically less profitable than are non-transactions (savings and time deposits) deposits.[28] Moreover, depositors who maintain a high balance are probably more lucrative. We do not have data on average account size. However, we can control for the type of deposit. The dependent variable in the bottom panel in Table 6 is the value of non-transactions deposits (savings and time deposits) held at the average

[26] Even if branches are relevant to producing small business loans, deposits held at a branch are likely to be positively correlated with the amount of small business loans associated with that branch because such loans involve the borrower opening a deposit account (and this account is likely to be booked to the branch involved in originating the loans).

[27] Deposit data are also from the FDIC's *Summary of Deposits* database. Only deposits held by individuals, partnerships and corporations are included. Government deposits are dropped because location is assumed to play only a small part if any in the holding of government deposits.

[28] Checking accounts are costly because processing check transactions is costly. These costs are thought to out-weigh the interest cost of savings and time deposits.

branch in a tract at time t for the period 1980-85. (Non-transactions deposits data are unavailable for the 1990s).

We find that the number of branches is negatively correlated with average branch deposits in relatively thinly-branched tracts (up to six branches in the 1990s and four branches in the 1980s). That is, an increase in the number of branches appears to be associated with *decreased* revenues. This effect is large; Table 6 suggests that adding a branch to a single branch tract over the 1990-95 period decreases total deposits by $25 million (which amounts to 16 percent of the unconditional mean total deposits at the typical branch in New York City of $155 million). We conclude that the apparent herding of bank branches is not due to unmeasured profitability differences across census tracts, at least not for relatively thinly branched tracts.

6. Interpretations and implications

The results reported in the previous section indicate that the number of existing branches in a tract has a positive, statistically significant effect on the number of subsequent openings and suggest, moreover, that this effect is not spurious. We interpret this as evidence of rational herding by banks, of the sort described in Section 3. But economic theory suggests at least two other categories of explanations for the clustering of firms, in which the number of firms located in an area directly influences the location decisions of subsequent firms.

The first set of potential explanations has to do with the presence of positive locational externalities.[29] These externalities may arise in a number of ways. For instance, the clustering of firms, by reducing consumers' search costs can increase aggregate demand. Dudey (1990) identifies conditions for an equilibrium where such clustering occurs as firms tradeoff the increased competition from locating close to competitors against the increased demand from such agglomeration. This explanation for the clustering of bank branches would be plausible if banking services were "search goods"—i.e., durable goods that are purchased fairly infrequently; that seems unlikely, however, given the survey findings mentioned in Section 2, which indicate that bank customers appear to value proximity to their bank mainly because the average bank customer visited a branch three times a month.[30]

Positive locational externalities might also arise if existing bank branches, through their lending operations, raise the deposit potential of a neighborhood. The problem with this explanation is that to the extent that banks enjoy first-comer advantages, individual banks ought to be able to internalize these dynamic externalities by expanding the scope of their operations within a neighborhood.

A second possible explanation for clustering is provided by models where in the presence of exogenous restrictions on price competition, firms compete for market share through locational choice.[31] Apart from the fact that such models yield clustering equilibria only for certain configurations of the spatial distribution of demand,

[29] As Devenow and Welch (1996) point out, the clustering of firms because of positive locational externalities (or in their terminology, payoff externalities) can also be considered a form of rational herding.

[30] Note though that the locational externalities that arise from reductions in travel costs might well explain why, *within* a neighborhood, retail businesses tend to locate along the main commercial thoroughfare.

[31] The first model of this kind appeared in Hotelling (1929).

configurations that seem unlikely to correspond to the distribution of demand for banking services in New York City, the assumption that locational choice is the only dimension along which banks compete seems unreasonable. While it is true that until the early 1980s banks were subject to strict interest rate regulation—which might be thought of as a restriction on price competition—many of these regulations were eased during the 80s, and even in the 1970s there is ample anecdotal evidence that banks competed through other means such as special promotional efforts.

These alternative explanations of branch clustering share a common prediction: clustering should increase branch profitability. That is, in terms of our notation, in the expression for branch-level profits:

$$\pi_{jt} = X_{jt}\alpha + \delta N_{jt} + e_{jt}$$

the coefficient on the number of existing branches should be positive, i.e., $\delta > 0$. To the extent that deposits are an adequate proxy for profits, the results reported in Table 6 reject this hypothesis—average branch-level deposits drop as the number of branches in a tract goes up, at least for the range of branches found in most tracts. Hence, we discount such explanations.

In contrast, the models of branch herding we discussed in Section 3 do predict that branch clustering can lower earnings. For example, information cascade models suggest that banks may herd and over-enter an area, driving down profits. Moreover, this effect can persist for two reasons: first, branch closures may be costly (and in Section 2 we provide some indirect evidence that this is the case); second, even if profits are lower in heavily-branched tracts, as long as banks continue to earn positive profits, there may be little incentive to explore the possibility that other, currently virgin tracts, may generate higher profits.

7. Conclusion

Bank branches in New York City tend to be spatially clustered. This unevenness has attracted considerable attention, both in the popular media and in policy circles because community groups argue that a bricks-and-mortar branch presence is important for access to banking services. In this paper we explored empirically whether the apparent clustering of bank branches can be at least partially attributed to rational herding by banks. We find that even after controlling for the expected profitability of operating a branch in an area, branch openings follow other, existing branches. Moreover, such bandwagon behavior appears to reduce branch profits. These findings, combined, suggest that herd behavior may be a factor in the branch location decisions of banks.

The primary implication of our finding that banks may be engaging in herd behavior is that the observed distribution of bank branches is potentially more skewed than the distribution of demographic and economic factors that affect branch profitability. Some neighborhoods may have an excessive number of branches while others remain underserved. In such a situation, there may be a possible governmental role in influencing the branch location decisions of banks. Unless the government is itself better informed than banks about the profit potential of different neighborhoods (which seems unlikely), the obvious policy instrument would be some form of subsidy that encourages experimentation, e.g., a subsidy to banks that open

branches in virgin territory. There may also be a more indirect role in encouraging the generation and dissemination of information about the characteristics of different neighborhoods.

References

[1] Allen, James (1995), "Branches ancient history? Not in this market," *American Banker*, February 13.

[2] Avery, Robert (1991), "Deregulation and the location of financial institution offices," *Economic Review*, Federal Reserve Bank of Cleveland, Third Quarter.

[3] Banking Policy Report (1994), "Justice Department Attacks Bank Marketing, Branching Patterns for First Time," September, v.5: 4-6

[4] Bikhchandani, Sushil, David Hirshleifer and Ivo Welch (1992), "A Theory of Fads, Fashion and Cultural Change as Information Cascades," *Journal of Political Economy*, October, v.100: 993-1026.

[5] Banerjee, Abhijit (1992), "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, August, v.CVII: 797-817.

[6] Calem, Paul (1995), "Mortgage Credit Availability in Low-and-Moderate-Income Minority Neighborhoods: Are Information Externalities Critical?" Working Paper No. 95-16, Board of Governors of the Federal Reserve System, Washington, D.C.

[7] Caplin, Andrew and John Leahy (1993), "Miracle on Sixth Avenue: Information Externalities and Search," Discussion Paper 681, Department of Economics, Columbia University.

[8] Caskey, John (1991), "Check-Cashing Outlets in the U.S. Financial System," *Economic Review*, Federal Reserve Bank of Kansas City, 76(6):53-67.

[9] Chamley, C. and D. Gale (1994), "Information revelation and strategic delay in a model of investment," *Econometrica*, September.

[10] DeCoster, Gregory and William Strange (1993), "Spurious agglomeration," *Journal of Urban Economics*, v.33: 273-304.

[11] Devenow, Andrea, and Ivo Welch (1996), "Rational herding in financial economics," *European Economic Review*, v.40, pp.603-615.

[12] Dudey, Marc (1990), "Competition by Choice: The Effect of Consumer Search on Firm Location Choice," *American Economic Review*, v.81: 1092-1104.

[13] Federal Deposit Insurance Corporation (1996), *Historical Statistics on Banking, 1934-1995*, Washington, D.C.

[14] Gale, Douglas (1996), "What Have We Learned from Social Learning?" *European Economic Review*, v.40: 617-628.

[15] Glassman, Cynthia (1995), "Industry role: who should serve the financial needs of 'underserved' communities?," *Journal of Retail Banking*, December.

[16] Grinblatt, Mark, Sheridan Titman and Russ Wermers (1995), "Momentum Investment Strategies, Portfolio Performance and Herding: A Study of Mutual Fund Behavior," *American Economic Review*, December, v85: 1088-1104.

[17] Gul, Faruk and Russell Lindholm (1995), "Endogenous timing and the clustering of agents' decisions," *Journal of Political Economy*, October.

[18] Hendricks, Kenneth and Dan Kovenock (1989), "Asymmetric information, information externalities, and efficiency: the case of oil exploration," *Rand Journal of Economics*, Summer.

[19] Hotelling, H. (1929), "Stability in competition," *Economic Journal*, 39: 41-57.

[20] Jain, Arvind and Satyadev Gupta (1987), "Some Evidence on 'Herding' Behavior of U.S. Banks," *Journal of Money, Credit and Banking*, February, v. 19: 78-89.

[21] Kutler, Jeffrey (1996), "Reports of the branch's demise have been greatly exaggerated," *American Banker*, December 31.

[22] Lang, William and Leonard Nakamura (1993), "A Model of Redlining," *Journal of Urban Economics*, v.33: 223-234.

[23] Lee, In Ho (1993), "On the convergence of informational cascades," *Journal of Economic Theory*, December.

[24] Loeb, P., W. Cohen, and C. Johnson (1995), "The new redlining: loan discrimination," *U.S. News and World Report*, April 17.

[25] Mead, Wendy (1997), "Banks: right mix of branches, alternative delivery is nearer," *American Banker*, March 12.

[26] Radecki, Lawrence, John Wenninger, and Daniel Orlow, (1996) "Bank Branches in Supermarkets," Federal Reserve Bank of New York, *Current Issues in Economics and Finance*, December, v.2.

[27] Scharfstein, David and Jeremy Stein (1990), "Herd Behavior and Investment," *American Economic Review*, June, v.80: 465-479.

[28] Tootell, Geoffrey (1996), "Redlining in Boston: Do Mortgage Lenders Discriminate Against Neighborhoods?," *Quarterly Journal of Economics*, November, pp.1049-1080.

[29] Welch, Ivo (1992), "Sequential sales, learning, and cascades," *Journal of Finance*, v.47, no.2, pp.695-732.

[30] Zweibel, Jeffrey (1995), "Corporate conservatism and relative compensation," *Journal of Political Economy*, v.103, no.1, 1-25.

Appendix

In this appendix we present a simple model of bank branch location which yields rational herding based on an informational externality along the lines of Lang and Nakamura (1993).

Suppose that the (true) profitability of operating a branch in tract j is given by:

$$\pi_j = X_j\beta - \delta(N_j + 1) + \upsilon_j$$

where X_j is a vector of observable characteristics of the tract, N_j is number of existing branches in the tract, $\delta \geq 0$ captures the possible (adverse) effect on profits from competition among branches in the tract, and υ_j is an unobserved tract-specific effect. We assume that there is an exogenously given probability, λ, that a bank, i, will consider opening a branch in tract j. Once a bank decides to explore the possibility of opening a branch in tract j, the bank receives a noisy private signal, ω_{ij}, say from a site analysis it commissions, as well as N_j noisy signals, $\omega_{kj}, k = 1, ..., N_j$, e.g., from the deposit levels of the N_j existing branches, about the profitability of operating in tract j. Let:

$$\omega_{kj} = \pi_j + \eta_{kj} \qquad\qquad \eta_{kj} \sim i.i.d\ N(0, \sigma_\eta^2)$$

The bank uses these signals to update its priors about the unobserved tract-effect, υ_j. We assume that the bank has unbiased priors regarding the value of υ_j; specifically we assume that the bank's prior μ_j is normally distributed with mean υ_j and variance σ_μ^2.

We assume that the bank is risk-averse and that its preferences can be represented by the exponential utility function:

$$U(\pi_j) = \frac{1}{\alpha} - \frac{1}{\alpha}e^{-\alpha\pi_j}$$

The bank therefore decides to open a branch in tract j only if:

$$E[U(\pi_j) \mid \omega_j; \mu_j] > 0$$

i.e., if its expected utility from opening a branch, given its prior, and given the signals it receives, is positive. Given the assumption of normality, and the exponential utility specification, this can be rewritten as:

$$E[U(\pi_j) \mid \omega_j; \mu_j] = \alpha E[\pi_j \mid \omega_j; \mu_j] - \frac{1}{2}\alpha^2 V[\pi_j \mid \omega_j; \mu_j] > 0$$

where $E[.]$ is the expectation and $V[.]$, the variance, of π_j, given ω_j and μ_j.

Now, the conjugacy property of the normal distribution implies that the bank's posterior beliefs about υ_j are also distributed normally with mean:

$$\upsilon_j + \gamma(N_j)\bar{\eta}_j$$

where:

$$\gamma(N_j) = \frac{(N_j + 1)\sigma_\mu^2}{(N_j + 1)\sigma_\mu^2 + \sigma_\eta^2} \text{ and } \bar{\eta}_j = \frac{1}{(N_j + 1)}\Sigma_k \eta_{kj}$$

Keeping in mind the assumption of unbiased priors, the expected profitability of operating in tract j can then be written:

$$E[\pi_j \mid \omega_j; \mu_j] = X_j\beta - \delta(N_j + 1) + \upsilon_j + \gamma(N_j)\bar{\eta}_j$$

The variance of π_j is given by:

$$V[\pi_j \mid \omega_j; \mu_j] = \frac{\sigma_\eta^2 \sigma_\mu^2}{(N_j + 1)\sigma_\mu^2 + \sigma_\eta^2}$$

Thus a bank will open a branch in tract j only if:

$$\alpha[X_j\beta - \delta(N_j + 1) + \upsilon_j] - \frac{1}{2}\alpha^2[\frac{\sigma_\eta^2 \sigma_\mu^2}{(N_j + 1)\sigma_\mu^2 + \sigma_\eta^2}] + \alpha\gamma(N_j)\bar{\eta}_j > 0$$

and the probability of this occurring is given by:

$$\Pr(\alpha\gamma(N_j)\bar{\eta}_j > -\alpha[X_j\beta - \delta(N_j + 1) + \upsilon_j] + \frac{1}{2}\alpha^2[\frac{\sigma_\eta^2 \sigma_\mu^2}{(N_j + 1)\sigma_\mu^2 + \sigma_\eta^2}])$$

In this expression the existing number of branches in a tract has two opposing effects on an entrant's branch opening decision. The first effect is the "competitive" effect captured by δ which should encourage banks to open branches in thinly branched tracts. The second effect, which does not appear in the "true" data generating process for profits, but appears as a result of the bank's preference for increased precision, tilts banks away from opening branches in thinly branched tracts. It is this second effect which raises the possibility of rational herding. Note also, that in this model, over-clustering due to herding is more likely to occur in the more profitable tracts since these tracts are the ones where there are likely to be more branches to begin with. Note also that the herding effect is strongest in thinly branched tracts. This is seen clearly in that the variance of expected profits $V[\pi_j|\omega_j; \mu_j]$ decreases at a decreasing rate in the number of branches, N_j. The reason is that adding a branch in a thinly branched area generates relatively more information about tract profitability than adding a branch to an area that already has many branches.

Table 1
Spatial distribution of bank branches in New York City
June 1990

No. of branches in tract: June 1990	Census tracts No.	Census tracts Frac.	Bank branches No.	Bank branches Frac.
0	1747	0.79	0	0.00
1	311	0.14	311	0.34
2	87	0.04	174	0.19
3	31	0.01	93	0.10
4	11	0.00	44	0.05
5	9	0.00	45	0.05
6-9	10	0.00	72	0.09
10-25	12	0.00	174	0.09
Total	2218	1.00	913	1.00

June 1995

No. of branches in tract: June 1995	Census tracts No.	Census tracts Frac.	Bank branches No.	Bank branches Frac.
0	1769	0.80	0	0.00
1	295	0.13	295	0.35
2	81	0.04	162	0.19
3	40	0.02	120	0.14
4	8	0.00	32	0.04
5	4	0.00	20	0.02
6-9	9	0.00	64	0.08
10-19	12	0.00	151	0.18
Total	2218	1.00	844	1.00

Note: Using the census definition, the following five counties make up New York city: Bronx, Kings, Queens, New York, and Richmond

Source: FDIC Summary of Deposits database, 1990-95. Covers full-service branches of all commercial banks.

Table 2
Spatial distribution of bank branch openings in New York City
July 1990 to June 1995

No. of branch openings in tract: July 1990–June 1995	Census tracts		Branch openings	
	No.	Frac.	No.	Frac.
0	2076	0.930	0	0.00
1	104	0.050	104	0.47
2	16	0.007	32	0.14
3	15	0.007	45	0.20
4	2	0.001	8	0.04
5	2	0.001	10	0.04
6-9	3	0.001	22	0.10
Total	2218	1.00	221	1.00

No. of branches in tract: June 1990	Census tracts		Branch openings	
	No.	Frac.	No.	Frac.
0	35	0.25	40	0.18
1	39	0.27	48	0.22
2	23	0.16	26	0.12
3	14	0.10	19	0.09
4	5	0.03	8	0.04
5	5	0.03	11	0.05
6-9	9	0.06	21	0.09
10-25	12	0.08	48	0.22
Total	142	1.00	221	1.00

Table 3
Summary statistics: New York City census tracts

Variable Means	All tracts	Branches: June1990 Tracts with branches	Tracts without branches
No. of tracts	2218	471	1747
Population of tract	3304	4361	3020
Fraction of population non-white	0.55	0.43	0.58
Fraction of population over age 65	0.13	0.15	0.13
Fraction of households poor	0.18	0.14	0.19
Fraction of population high-school graduates	0.67	0.73	0.66
Median household income ($)	37043	45663	34728
Fraction of housing units rental	0.65	0.69	0.64
Median value of owner-occupied housing	164246	184130	158908
Number of people working in tract	634	1715	333
Indicator that tract is net-importer of workers	0.24	0.46	0.18
Fraction of land area commercial	0.06	0.13	0.04
Fraction of land area industrial	0.05	0.05	0.05
Fraction of land area single-family residences	0.36	0.26	0.38
Fraction of land area multi-family residences	0.23	0.25	0.22
No. of existing branches: June 1990	0.41	1.94	0.00
No. of branch openings: July 1990–June 1995	0.10	0.46	0.00

Variable	All tracts	Branch openings: July 1990–June 1995 Tracts with branch openings	Tracts without branch openings
No. of tracts	2218	142	2076
Population of tract	3304	4728	3207
Fraction of population non-white	0.55	0.37	0.56
Fraction of population over age 65	0.13	0.15	0.13
Fraction of households poor	0.18	0.14	0.18
Fraction of population high-school graduates	0.67	0.76	0.67
Median household income ($)	37043	53602	35917
Fraction of housing units rental	0.65	0.72	0.64
Median value of owner-occupied housing	164246	175690	163468
Number of people working in tract	639	3692	421
Indicator that tract is net-importer of workers	0.24	0.60	0.21
Fraction of land area commercial	0.06	0.21	0.04
Fraction of land area industrial	0.05	0.04	0.05
Fraction of land area single-family residences	0.36	0.16	0.37
Fraction of land area multi-family residences	0.23	0.27	0.22
No. of existing branches: June 1990	0.41	2.90	0.24
No. of branch openings: July 1990–June 1995	0.10	1.55	0.00

Table 4
Evidence on herding: basic results

Dependent variable: no. of branch openings, 1990-95	Poisson	Ordered logit
No. of branches	0.32	0.41
	(5.1)	(5.4)
No. of branches squared	-0.01	-0.01
	(-4.3)	(-3.7)
Population of tract	9.9×10^{-5}	1.0×10^{-4}
	(2.6)	(2.5)
Fraction of population non-white	-0.83	-1.04
	(-1.77)	(-2.0)
Fraction of population over age 65	1.0	1.4
	(0.94)	(1.1)
Fraction of households poor	-3.1	-2.4
	(-2.4)	(-1.6)
Fraction of population high-school graduates	-1.7	-1.0
	(-1.58)	(-0.83)
Median household income ($)	-8.1×10^{-6}	-1.1×10^{-5}
	(-0.76)	(-0.9)
Fraction of housing units rental	-0.32	-0.54
	(-0.5)	(-0.65)
Median value of owner-occupied housing	2.5×10^{-7}	-1.7×10^{-7}
	(-0.43)	(-0.26)
Number of people working in tract	-2.5×10^{-5}	-1.1×10^{-5}
	(0.68)	(-0.25)
Indicator that tract is net-importer of workers	0.66	0.61
	(2.9)	(2.6)
Fraction of land area commercial	0.26	0.032
	(3.3)	(3.6)
Fraction of land area industrial	-0.01	-0.02
	(-1.1)	(-1.5)
Fraction of land area single-family residences	-0.02	-0.017
	(-2.4)	(-2.2)
Fraction of land area multi-family residences	0.01	0.01
	(1.3)	(1.3)
No. of observations	10,075	10,075
Pseudo-R-square	0.32	0.28

Notes: (1) t-statistics appear below coefficient estimates. (2) Regressions based on pooled annual tract-level data for New York City from 1990–95. (3) Regressions include changes in tract-level demographic variables between 1980 and 1990; coefficient estimates are not reported.

Table 5
Summary of regressions of branch openings on tract features

Dependent variable (sample)	No. of branches	No. of branches squared	No. of obs.	Pseudo R-squared
No. of openings, annnually, of all branches: 1990-95				
Poisson	0.32	-0.01	10075	0.31
	(5.1)	(-4.3)		
Ordered logit	0.41	-0.01	10075	0.27
	(5.4)	(-3.7)		
Cumulative no. of openings of all branches: 1990-95				
Poisson	0.31	-0.01	2015	0.43
	(5.1)	(-4.3)		
Ordered logit	0.58	-0.01	2015	0.33
	(4.9)	(-2.2)		
No. of openings, annually, of branches other than headoffices: 1992-95				
Poisson	0.44	-0.02	6045	0.27
	(5.1)	(-4.3)		
Ordered logit	0.55	-0.02	6045	0.24
	(5.1)	(-3.6)		
Cumulative no. of openings of branches other than headoffices: 1992-95				
Poisson	0.43	-0.02	2015	0.33
	(4.9)	(-4.4)		
Ordered logit	0.63	-0.02	2015	0.28
	(4.6)	(-3.1)		
No. of openings, annnually, of all branches: 1980-85				
Poisson	0.65	-0.02	10570	0.31
	(9.5)	(-5.7)		
Ordered logit	0.69	-0.02	10570	0.27
	(8.4)	(-4.2)		
Cumulative no. of openings of all branches: 1980-85				
Poisson	0.67	-0.02	2114	0.40
	(8.9)	(-5.0)		
Ordered logit	0.79	-0.02	2114	0.27
	(5.9)	(-1.4)		

Notes: (1) t-statistics appear below coefficient estimates; (2) The regression model used for the 1990-95 data is the same as that used in Table 4. The model used for the 1980-85 data includes the following tract characteristics as controls: population, race, poverty rate, age, rental rate, median income and median rent. (3) "Cumulative openings" equals the total number of branches that opened in a tract over the relevant period. (4) Reliable data on head offices are available only from 1992.

Table 6
Further tests: Summary of regressions of average branch deposits on tract features

Dependent variable:	Number of branches (OLS)	Number of branches squared (OLS)	N	R-squared
Average total deposits per branch				
Annual: 1990-95	-29,501	2381	2589	0.41
	(-6.6)	(7.4)		
Annual: 1984-85	-29,491	3310	954	0.58
	(-4.2)	(5.3)		
Average non-transactions deposits per branch				
Annual: 1980-85	-21,267	2280	2961	0.31
	(-5.2)	(6.1)		

Notes: (1) t-statistics appear below coefficient estimates. These t-statistics are based on heteroskedasticity-consistent standard errors [White 1980]. (2) The regression models used are the same as those used in Table 5. But the dependent variable now is the average deposit level across all the branches in a tract in a given year. Total deposits consist of all deposits held by everyone other than the government. Non-transactions deposits consist of savings and time deposits held by everyone other than the government. (3) Reliable data on total non-government deposits are available only from 1984. Non-transactions data are not available for the 1990s.(4) Sample means are as follows: Mean total deposits at the typical branch between 1990 and 1995 was $155 million; the mean for 1984-85 was $110 million; mean non-transactions deposits for 1980-85 was $62 million.

FEDERAL RESERVE BANK OF NEW YORK
RESEARCH PAPERS
1997

The following papers were written by economists at the Federal Reserve Bank of New York either alone or in collaboration with outside economists. Single copies of up to six papers are available upon request from the Public Information Department, Federal Reserve Bank of New York, 33 Liberty Street, New York, NY 10045-0001 (212) 720-6134.

9701. Chakravarty, Sugato, and Asani Sarkar. "Traders' Broker Choice, Market Liquidity, and Market Structure." January 1997.

9702. Park, Sangkyun. "Option Value of Credit Lines as an Explanation of High Credit Card Rates." February 1997.

9703. Antzoulatos, Angelos. "On the Determinants and Resilience of Bond Flows to LDCs, 1990 - 1995: Evidence from Argentina, Brazil, and Mexico." February 1997.

9704. Higgins, Matthew, and Carol Osler. "Asset Market Hangovers and Economic Growth." February 1997.

9705. Chakravarty, Sugato, and Asani Sarkar. "Can Competition between Brokers Mitigate Agency Conflicts with Their Customers?" February 1997.

9706. Fleming, Michael, and Eli Remolona. "What Moves the Bond Market?" February 1997.

9707. Laubach, Thomas, and Adam Posen. "Disciplined Discretion: The German and Swiss Monetary Targeting Frameworks in Operation." March 1997.

9708. Bram, Jason, and Sydney Ludvigson. "Does Consumer Confidence Forecast Household Expenditure: A Sentiment Index Horse Race." March 1997.

9709. Demsetz, Rebecca, Marc Saidenberg, and Philip Strahan. "Agency Problems and Risk Taking at Banks." March 1997.

9710. Lopez, Jose. "Regulatory Evaluation of Value-at-Risk Models." March 1997.

9711. Cantor, Richard, Frank Packer, and Kevin Cole. "Split Ratings and the Pricing of Credit Risk." March 1997.

9712. Ludvigson, Sydney, and Christina Paxson. "Approximation Bias in Linearized Euler Equations." March 1997.

9713. Chakravarty, Sugato, Sarkar, Asani, and Lifan Wu. "Estimating the Adverse Selection Cost in Markets with Multiple Informed Traders." April 1997.

9714. Laubach, Thomas, and Adam Posen. "Some Comparative Evidence on the Effectiveness of Inflation Targeting." April 1997.

9715. Chakravarty, Sugato, and Asani Sarkar. "A General Model of Brokers' Trading, with Applications to Order Flow Internalization, Insider Trading and Off-Exchange Block Sales." May 1997.

9716. Estrella, Arturo. "A New Measure of Fit for Equations with Dichotomous Dependent Variables." May 1997.

9717. Estrella, Arturo. "Why Do Interest Rates Predict Macro Outcomes? A Unified Theory of Inflation, Output, Interest and Policy." May 1997.

9718. Ishii Jun, and Kei-Mu Yi. "The Growth of World Trade." May 1997.

9719. Kambhu, John. "Interest Rate Options Dealers' Hedging in the US Dollar Fixed Income Market." May 1997.

9720. Kroszner, Randall, and Philip Strahan. "The Political Economy of Deregulation Evidence from the Relaxation of Bank Branching Restrictions in the United States." June 1997.

9721. Locke, Peter, Sarkar, Asani, and Lifan Wu. "Market Liquidity and Trader Welfare in Multiple Dealer Markets: Evidence from Dual Trading Restrictions." July 1997.

9722. Zakrajšek, Egon. "Retail Inventories, Internal Finance, and Aggregate Fluctuations: Evidence from U.S. Firm-Level Data." July 1997.

9723. Lown, Cara, and Robert Rich. "Is There An Inflation Puzzle?" August 1997.

To obtain more information about the Bank's *Research Papers* series and other publications and papers, visit our site on the World Wide Web (http://www.ny.frb.org). From the research publications page, you can view abstracts for *Research Papers* and *Staff Reports* and order the full-length, hard copy versions of them electronically. Interested readers can also view, download, and print any edition in the *Current Issues in Economics and Finance* series, as well as articles from the *Economic Policy Review.*