Working Paper Series

# Internet Banking: An Exploration in Technology Diffusion and Impact

WP 13-10R

Richard Sullivan
Federal Reserve Bank of Kansas City

Zhu Wang
Federal Reserve Bank of Richmond



THE FEDERAL RESERVE BANK OF RICHMOND
RICHMOND ▪ BALTIMORE ▪ CHARLOTTE

# Internet Banking: An Exploration in Technology Diffusion and Impact*

Richard Sullivan†, Zhu Wang‡

November 2017
Working Paper No. 13-10R

**Abstract**

Taking Internet banking as an example, we study diffusion of cost-saving technological innovations and the impact on firm size distribution. In doing so, we construct a competitive banking industry model where banks differ in size due to cost heterogeneity. The model matches the actual bank size distribution well and generates $S$-shaped logistic diffusion curves as documented in the literature. We apply the theory to an empirical study of Internet banking diffusion among banks across 50 U.S. states. Our findings disentangle the interrelationship between Internet banking diffusion and bank size distribution, and explain the variation in diffusion rates across geographic regions.

*JEL classification:* G20; L10; O30

*Keywords:* Technology diffusion; Bank size distribution; Internet banking

*We thank Dean Amel, Boyan Jovanovic, Ned Prescott and participants at the NBER Summer Institute Productivity Potpourri Meeting, Federal Reserve System Applied Microeconomics Conference, Federal Reserve System Financial Structure and Regulation Meeting, and various seminars for helpful comments. Nathan Halmrast, Christian Hung and Emily Cuddy provided valuable research assistance. The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Kansas City, Federal Reserve Bank of Richmond, or the Federal Reserve System.

†Federal Reserve Bank of Kansas City; rick.j.sullivan@kc.frb.org.

‡Federal Reserve Bank of Richmond; zhu.wang@rich.frb.org.

# 1 Introduction

New ideas, embodied in product and technology innovations, are fundamental driving forces for long-run growth. However, it often takes a lengthy period of many years for an innovation to become widely adopted by the population, a process termed as "diffusion." Moreover, the speed of diffusion is rarely constant. Rather, we typically observe $S$-shaped diffusion curves that resembles a logistic function. To better understand the diffusion process, an extensive literature has been developed that seeks to explain how, why, and at what rate new ideas and technology spread.

A majority of literature emphasizes the role played by communication of information (Rogers, 2003). One most popular theory focuses on *contagion*, or so-called "word-of-mouth" effect, that agents adopt innovations when they come in contact with others who have already adopted; in other words, innovations spread like epidemics. Two alternative but related theories are *social influence* and *social learning*, which attribute contagion to social forces such as conformity motive or belief updating. A common theme of these theories is that the diffusion process is driven by internal feedback effects from prior to future adopters (see, e.g. Young 2009 for an overview of the "internal diffusion" models). These models are particularly appealing for empirical uses because the internal feedback effect can be formalized as a differential equation that generates logistic diffusion curves (e.g., Griliches 1957, Mansfield 1961, Bass 1969, 2004).

In contrast, a competing view in the literature emphasizes agents' heterogeneity in terms of adoption costs and benefits (e.g., David 1969, 2005, Stoneman 2002). According to that view, diffusion lags are not necessarily explained by incomplete information. Rather, agents may have complete information and make adoption decisions based on their heterogenous willingness to pay for the innovation. As a result, diffusion is driven mainly by *external* factors, such as price and quality changes, and diffusion curves can be $S$-shaped if the adoption thresholds of agents follow a positively skewed distribution.

In this paper, we incorporate and extend the ideas from the literature and study the diffusion of a recent technological innovation, Internet banking. Our study goes beyond the diffusion process *per se* to also explore the reverse effect of diffusion on industry development, particularly the shift of firm size distribution (or, bank size distribution in the Internet banking context). In doing so, we first construct a competitive banking industry model where banks differ in size due to cost heterogeneity, and the model-implied bank size distribution well matches the actual distribution. Our theory suggests that as Internet banking is initially introduced, large banks enjoy cost advantages in adopting it early on and thus increase their size relative to non-adopters. Over time, due to external changes (e.g., demand shift, technological progress, and/or industry deregulation), the innovation gradually diffuses into smaller banks. This approach is consistent with the *external diffusion* view mentioned above but also generates a closed-form logistic diffusion curve that resembles those derived from the *internal diffusion* models.

We then apply the theory to an empirical study of Internet banking diffusion among banks across 50 U.S. states. We show that our theory provides a parsimonious empirical framework. Particularly, the model implies estimating a simultaneous equation system, which jointly determines Internet banking diffusion and bank size distribution. We augment this equation system with empirical variables that control for technological, economic, and institutional factors as well as the potential contagion effect suggested by the *internal diffusion* models. Employing instrument variables in our simultaneous-equation estimation, we are able to disentangle the positive interactions between Internet banking diffusion and bank size distribution, and explain the variation in diffusion rates across U.S. geographic regions.

As mentioned above, our paper is directly related to the literature on technology diffusion and bridges a gap between the internal and external diffusion models. In the banking context, several recent studies have looked at the Internet and related technology adoption in the banking industry.[1] However, unlike our paper, those studies focus more

[1] For example, Hernández-Murillo et al. (2010) study a panel of commercial banks for 2003-2006 and show that banks adopt online banking earlier in markets where their competitors have already done so. DeYoung et al. (2007) study a sample of U.S. banks in the late 1990s. They find that branching intensity and online banking are complementary and online banking adoption positively affects the bank's future performance. Courchane et al. (2002) develop and estimate a model for early adoption of Internet banking. They find that relative bank size and demographic information predictive of future demand positively influence Internet banking adoption. Furst et al. (2001) estimate a logit model for Internet banking adoption in a sample of national banks. They find that larger banks and banks that are younger and better performing are more likely to adopt Internet banking.

on individual banks' adoption decision rather than the aggregate pattern of diffusion and bank size distribution.[2] Our paper is also naturally connected with the large and growing literature on industry dynamics and firm size distribution (e.g., Lucas 1979, Jovanovic 1982, Hopenhayn 1992, Sutton 1997, Cabral and Mata 2003, among many others). In the banking field, some recent studies have explored the size distribution of banks (e.g., Berger, Kashyap, and Scalise 1995, Ennis 2001, Jones and Critchfield 2005, Janicki and Prescott 2006, McCord and Prescott 2014). However, those studies are primarily interested in the effects of laws and regulations on bank size distribution, while the focus of this paper is technology diffusion.

The paper is organized as follows. Section 2 introduces industry background regarding the banking sector and Internet banking diffusion. Section 3 presents a competitive equilibrium model of technology diffusion and industry evolution in the Internet banking context. Section 4 applies the theory to an empirical study on Internet banking diffusion among banks across 50 U.S. states between 2003-2007. Section 5 concludes.

# 2 Industry background

In our study, Internet banking is defined as a bank providing a website that allows customers to execute transactions on their accounts. In the United States, the history of Internet banking can be traced back to 1995 when Wells Fargo first allowed its customers to access account balances online.[3] Ever since then, banks have steadily increased their online presence. Figure 1 plots the diffusion of Internet banking among in-state banks between 2003-2007, before the start of the Great Recession.[4] In-state banks refer to commercial banks focusing on operating in a single state, which accounted for more than 90

[2] Note that adoption and diffusion are two related but different terms used in the literature: Adoption typically refers to an individual process of adopting an innovation, while diffusion is a group phenomenon that refers to how an innovation spreads.

[3] Internet-only banks account for a very small fraction of the U.S. banking population (less than 0.5 percent even during the dot-com boom years). In this paper, we focus on the Internet banking adoption among traditional brick-and-mortar banks. See Wang (2007) for an analysis of Internet-only banks.

[4] Data Source: Call Report. Since 2003, depository institutions have been required to report whether their websites allow customers to execute transactions on their accounts. Our sample ends in 2007 because the adoption had become almost universal by then and we also want to avoid the disruption of the Great Recession.

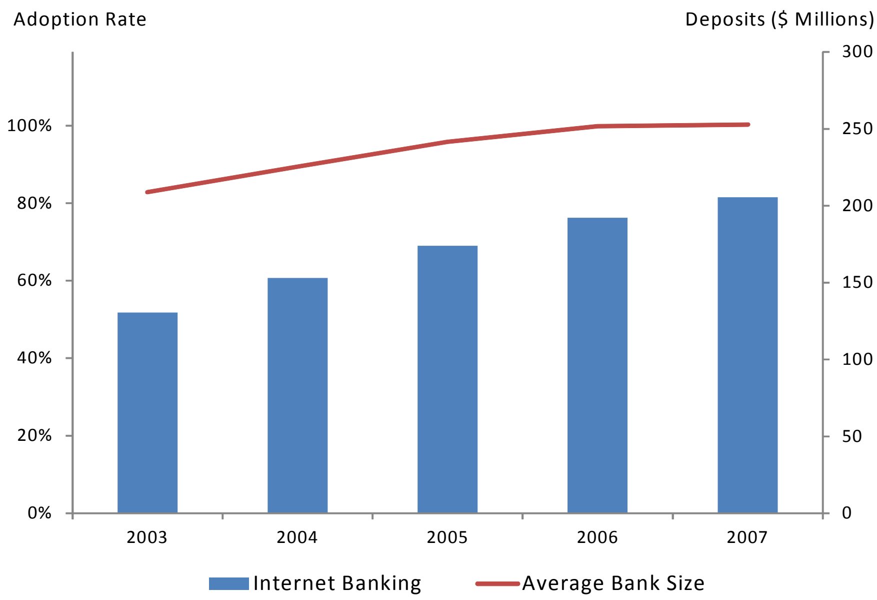


Figure 1: Internet Banking Diffusion and Average Bank Size

percent of the U.S. banking population during this period.[5] The figure shows that 51.8 percent of in-state banks had adopted Internet banking by 2003, and the ratio continued to rise to 81.5 percent in 2007.[6]

However, the diffusion pattern varies significantly across bank size groups and geographic regions. First, looking across size groups, large banks appear to have an advantage adopting the innovation than smaller ones. As shown in Figure 2, 90.5 percent of in-state banks with deposits over $300 million reported that they had a transactional website in 2003, compared to only 10.5 percent of in-state banks with deposits under $25 million. The variation is also striking across geographic regions. Figure 3 compares Internet banking adoption by in-state banks across U.S. states in 2003. The northeast and the west regions

[5] More specifically, a bank is classified as an in-state bank if all its deposits are in the state of the bank's headquarter. As will become clear, focusing on in-state banks allows us to avoid the complications of interstate banking when comparing Internet banking adoption and bank size distributions across states. In 2003, there were 7,712 commercial banks in the United States, among which 7,183 were in-state banks (i.e., 93 percent).

[6] A similar diffusion pattern can be found if we instead consider all U.S. commercial banks. By 2003, 53 percent of all commercial banks had adopted transactional websites, and the ratio rose to 82 percent in 2007. In the meantime, the number of U.S. households that were using Internet banking rose from 30 million in 2003 to 45 million in 2007 (Source: *Online Banking Report* #224, January 2014).

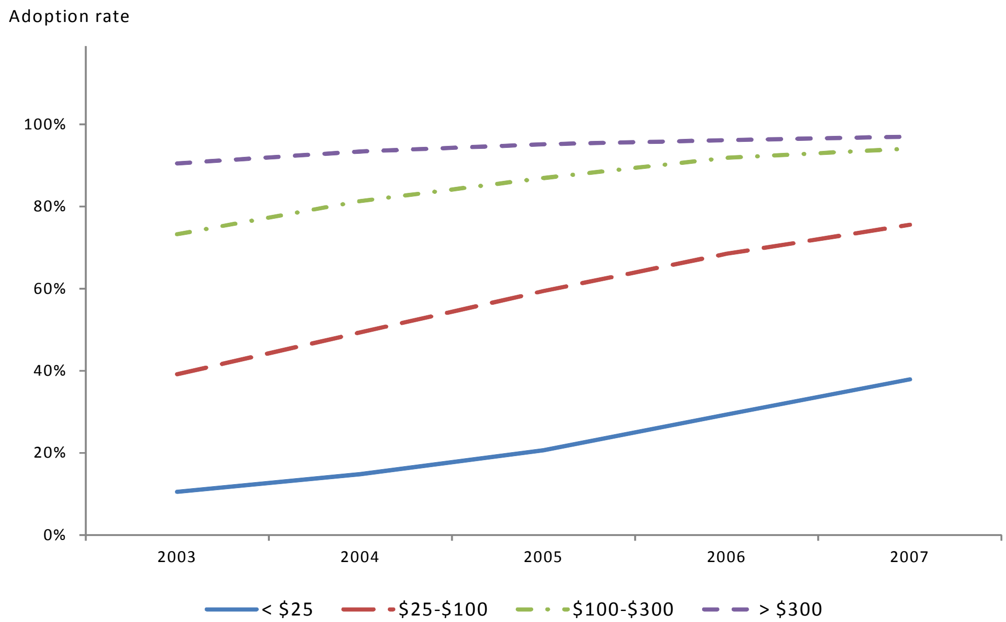


Figure 2: Internet Banking Adoption by Bank Size Group (Deposits: Millions)

had the highest adoption rates (i.e., 65-85 percent in each state), while the central regions of the country had the lowest (i.e., 25-45 percent in each state). These observations raise important questions regarding technology diffusion: Why do large banks tend to be early adopters of the Internet innovation? What determines the different diffusion rates across bank groups and geographic regions?

Meanwhile, the diffusion of Internet banking was accompanied with continuing changes of bank size distribution. The U.S. banking industry has been through major deregulation and consolidation since mid-1990s.[7] The number of commercial banks dropped substantially while the bank size distribution continue to shift (Figure 1 plots the average deposits of in-state banks between 2003-2007). Therefore, it is interesting to explore the role of Internet banking played in this process: Considering that bank size is an important factor for adopting Internet banking, how much has banking deregulation and consolidation affected Internet banking diffusion? At the same time, how much, if any, has Internet banking diffusion influenced the bank size distribution?

[7] According to *FDIC Quarterly Banking Profile Graph Book,* there were about 100 interstate bank mergers and 200 intrastate bank mergers per year between 2003 and 2007.

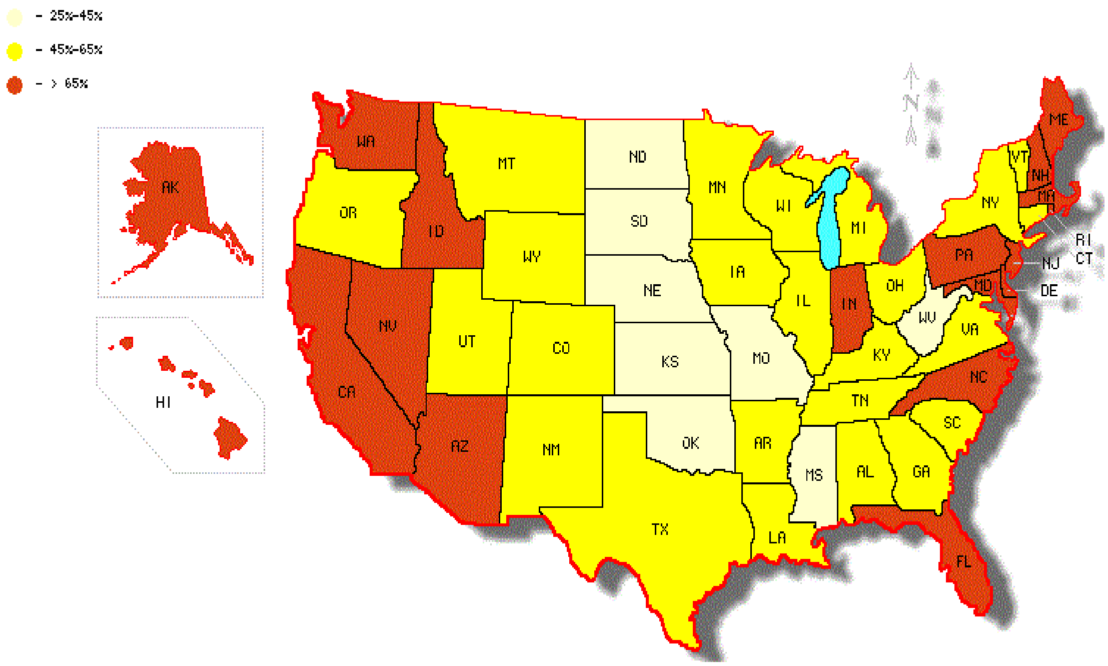


Figure 3: Internet Banking Adoption by State (2003)

These observations and questions motivate our study of Internet banking diffusion and impact. Conceptually, the benefits of Internet banking can be viewed as twofold. First, it brings convenience to bank customers, allowing them to use services from banks in distance and avoid hassles to go to ATMs or branches. Second, it generates substantial cost savings to banks. Most banking websites provide balance transfer and bill payments services, and some also process applications for deposits, loans and credit cards.[8] This allows banks to conduct standardized, low-value-added transactions through the online channel, while focusing their resources on more specialized, high-value-added transactions (e.g., business lending, personal trust services, investment banking) through branches. In fact, the ratio of bank employees (and bank tellers) to deposits have been declining since the late 1990s.[9] This is consistent with continuous progress in IT technology, including

[8]For instance, a survey conducted by the Federal Reserve Bank of Kansas City shows that in the tenth Federal Reserve District, more than 70 percent commercial bank websites provided balance transfer and bill payment services, and less than 20 percent allowed for online application for deposits, loans or credit cards in 2006.

[9]Between 1997 and 2007, the number of bank employees per million-dollar deposits fell from 0.44 to 0.24, and the number of bank tellers per million-dollar deposits fell from 0.14 to 0.09. (Data sources: Commercial bank employees and tellers are from the BLS, and commercial bank deposits are from the

the increasing adoption of Internet banking.

Based on the facts and reasonings above, we will first construct a model in the following section that focuses on banks' cost saving motive of adopting Internet banking. The model considers a competitive banking industry, where banks' sizes are primarily determined by cost components (For simplicity, we abstract from consumers' convenience benefits of using Internet banking and banks' strategic motives of adoption in the theory, but those factors will later be incorporated in the empirical analysis).[10] We show that the model generates $S$-shaped logistic diffusion curves and allows us to trace industry dynamics including the shifting bank size distribution. We will then apply the theory to an empirical study of Internet banking diffusion among in-state banks across 50 U.S. states between 2003-2007.

# 3 The model

In this section, we construct a theoretical model of technology diffusion and firm size distribution. While the model is in the context of Internet banking, its implications are general and could be applied to cost-saving technological innovations in other industries.

## 3.1 Environment

The industry is composed of a continuum of banks which produce homogenous banking services. Banks behave competitively, taking the market price of banking services as given. We assume banks are heterogenous in productivity, which yields size differences. At a point in time $t$, the market demand takes a simple inelastic form – consumers are willing to pay $P_t$ for an amount $Y_t$ of banking services. Over time, $P_t$ and $Y_t$ might be shifted by economic forces, such as changes in population, consumer income, or competing services.[11]

FDIC).

[10] Alternatively, we could model a differentiated banking market, where banks engage in strategic competition on price and service levels. Such a model might be more realistic, but on the other hand could be too complicated to explain the high-level patterns of Internet banking diffusion and impact.

[11] Our empirical study will focus on in-state banks, a subsample of the banking population. Therefore, it is reasonable to assume that these (in-state) banks face exogenous $P$ and $Q$, which are determined by the overall banking market conditions, including the competition from large interstate banks. In fact, in the empirical study, we will include the out-of-state bank presence in the in-state banking market as a

## 3.2 Pre-innovation equilibrium

Before the technological innovation arrives, the industry is at a steady state. Taking the market price as given, an individual bank maximizes its profit under the existing technology:

$$\pi = \underset{y}{Max}\ Py - \alpha y^{\beta}$$

where $\pi$ is the bank's profit, $P$ is the price, $y$ is the bank's output, and $\alpha > 0$ and $\beta > 1$ are cost parameters.

Profit maximization yields

$$y = (\frac{P}{\alpha\beta})^{\frac{1}{\beta-1}}. \tag{1}$$

Banks are heterogenous in the cost parameter $\alpha$, so there is a distribution $G$ of bank size measured by output. Historically, bank size $y$ fits well with the log-logistic distribution (See Figure 4 for an example)[12], which has the cdf function

$$\Pr(y \leq x) = G_y(x) = 1 - \frac{1}{1 + b_1 x^{b_2}} \tag{2}$$

with the mean $E(y)$ and Gini coefficient $g$ given as

$$E(y) = b_1^{-1/b_2}\Gamma(1 + \frac{1}{b_2})\Gamma(1 - \frac{1}{b_2}), \qquad g = \frac{1}{b_2},$$

where $\Gamma$ denotes the gamma function $\Gamma(\mu) \equiv \int_0^\infty s^{\mu-1}\exp(-s)ds$.

Rewriting the log-logistic distribution into a more intuitive form, we have

$$G_y(x) = 1 - \frac{1}{1 + (\eta x/E(y))^{1/g}}, \tag{3}$$

where $\eta = \Gamma(1 + g)\Gamma(1 - g)$.

regressor to control for the demand for the services of in-state banks.

[12]Figure 4 uses deposits as a measure of bank size. We also used assets as an alternative measure of bank size and the plot is very similar. The log-logistic distribution is an easily tractable representative of the larger group of positively skewed distributions. As will become clear, it also connects our study to the typically observed logistic diffusion curves.

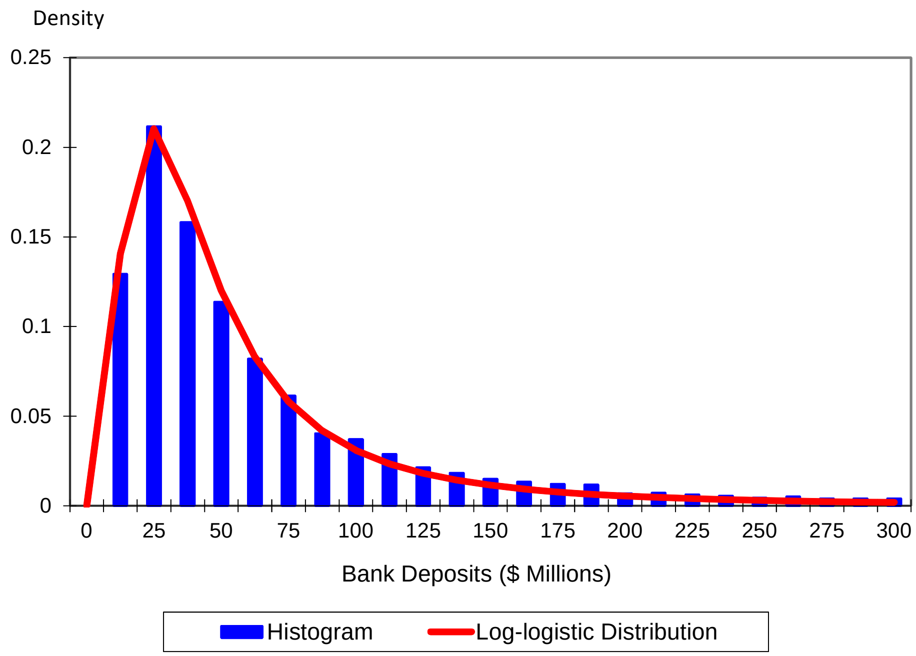


Figure 4: Bank Size Distribution (In-State Banks 1990)

At equilibrium, industry demand equals supply, so that

$$E(y)N = Y,$$

where $N$ is the number of banks.

Note that our assumption of log-logistic size distribution is robust to environmental changes. For example, shocks to price $P$, mean productivity $E(\alpha^{\frac{1}{1-\beta}})$, or demand $Y$ may affect the mean bank size $E(y)$ and/or the number of banks $N$, but not other properties of the distribution.[13]

## 3.3 Post-innovation equilibrium

### 3.3.1 Individual adoption decision

The technological innovation, Internet banking, arrives at a point in time (which we normalize as time 0). Thereafter, at each period, an individual bank decides whether to

[13] Note that $\alpha^{\frac{1}{1-\beta}}$ decreases in $\alpha$ for $\beta > 1$. Hence, $\alpha^{\frac{1}{1-\beta}}$ can be interpreted as a productivity measure.

adopt the innovation or not ($a$ = adopt; $n$ = not adopt):

$$\pi = Max\{\pi_n, \pi_a\},$$

$$\text{where} \quad \pi_n = \underset{y_n}{Max} Py_n - \alpha y_n^\beta, \qquad \pi_a = \underset{y_a}{Max} Py_a - \frac{\alpha}{\gamma} y_a^\beta - k.$$

Note that $\gamma > 1$ is the cost saving by adopting the innovation, and $k > 0$ is the period cost of adoption.[14]

Solving the maximization problems yields

$$\begin{aligned} y_n &= (\frac{P}{\alpha\beta})^{\frac{1}{\beta-1}} \quad, \quad \pi_n = \frac{\beta-1}{\beta} Py_n; \\ y_a &= (\frac{\gamma P}{\alpha\beta})^{\frac{1}{\beta-1}} \quad, \quad \pi_a = \frac{\beta-1}{\beta} Py_a - k. \end{aligned}$$

An individual bank adopts Internet banking iff $\pi_a \geq \pi_n$, and hence there is a threshold size $y_n^*$ for adoption:

$$\pi_a = \pi_n \Longrightarrow y_n^* = \frac{k}{(\frac{\beta-1}{\beta})(\gamma^{\frac{1}{\beta-1}} - 1)P}.$$

The size threshold for adoption suggests that large banks have an advantage adopting the innovation. Considering the randomness of environment in reality, this result is expected to hold statistically in the data, as shown in Figure 2.

#### 3.3.2 Aggregate adoption

Given the bank size distribution $G$ defined in Eq (3) and the adoption threshold $y_n^*$, the aggregate adoption rate of Internet banking is

$$F = 1 - G_{y_n}(y_n^*) = \frac{1}{1 + (\eta y_n^*/E(y_n))^{1/g}}, \tag{4}$$

$$\text{where} \qquad y_n = (\frac{P}{\alpha\beta})^{\frac{1}{\beta-1}}, \qquad y_n^* = \frac{k}{(\frac{\beta-1}{\beta})(\gamma^{\frac{1}{\beta-1}} - 1)P}.$$

We then derive the following Proposition 1.

[14] The period cost $k$ may include the rental cost of equipment and the cost of operating the website.

**Proposition 1** *The adoption rate $F$ increases in consumer willingness-to-pay $P$, mean bank productivity $E(\alpha^{\frac{1}{1-\beta}})$, and cost saving $\gamma$, but decreases in adoption cost $k$.*

**Proof.** Equation 4 implies that $\partial F/\partial P > 0$, $\partial F/\partial E(\alpha^{\frac{1}{1-\beta}}) > 0$, $\partial F/\partial \gamma > 0$, and $\partial F/\partial k < 0$. ■

#### 3.3.3 Average bank size

Note that $E(y_n)$ is not directly observable after Internet banking is introduced. The observed average bank size is

$$E(y) = \int_0^{y_n^*} y_n dG(y_n) + \int_{y_n^*}^{\infty} y_a dG(y_n) = E(y_n) + [\gamma^{\frac{1}{\beta-1}} - 1] \int_{y_n^*}^{\infty} y_n dG(y_n).$$

Given that $y_n$ follows the log-logistic distribution $G$, we have

$$\int_{y_n^*}^{\infty} y_n dG(y_n) = E(y_n)[1 - \beta(1+g, 1-g; G(y_n^*))],$$

where $\beta$ is the incomplete beta function defined as

$$\beta(a, b; x) \equiv \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^x s^{a-1}(1-s)^{b-1} ds \quad \text{with} \quad a > 0, b > 0, x > 0,$$

$$\beta(a, b; 0) = 0 \qquad \text{and} \qquad \beta(a, b; 1) = 1.$$

Therefore, the observed average bank size $E(y)$ can be derived as

$$E(y) = E(y_n)\{1 + [\gamma^{\frac{1}{\beta-1}} - 1][1 - \beta(1+g, 1-g; 1-F)]\}. \tag{5}$$

Proposition 2 then follows.

**Proposition 2** *The average bank size $E(y)$ increases in consumer willingness-to-pay $P$, mean bank productivity $E(\alpha^{\frac{1}{1-\beta}})$, and cost saving $\gamma$, but decreases in adoption cost $k$.*

**Proof.** Given Proposition 1, Eq (5) implies that $\partial E(y)/\partial P > 0$, $\partial E(y)/\partial \gamma > 0$, $\partial E(y)/\partial E(\alpha^{\frac{1}{1-\beta}}) > 0$ and $\partial E(y)/\partial k < 0$. ■

### 3.4 Industry dynamics

Equations (4) and (5) describe the post-innovation industry equilibrium at a point in time. Note that we have so far omitted time subscripts on all variables. To discuss the industry dynamics, we now add them back and show that the diffusion path derived from our model closely follows a logistic curve, a path well documented in the literature on technology diffusion.

Consider a banking industry under continuous environmental changes (e.g., demand shift, technological progress, and/or industry deregulation). As a result, consumer willingness-to-pay $P_t$, mean bank productivity $E(\alpha_t^{\frac{1}{1-\beta}})$, Internet banking cost saving $\gamma_t$, and adoption cost $k_t$ may change constantly. Therefore, we specify simple laws of motion as follows:

$$\begin{aligned} P_t &= P_0 e^{z_p t}, \quad \gamma_t^{\frac{1}{\beta-1}} - 1 = (\gamma_0^{\frac{1}{\beta-1}} - 1)e^{z_\gamma t}, \\ k_t &= k_0 e^{z_k t}, \quad E(\alpha_t^{\frac{1}{1-\beta}}) = E(\alpha_0^{\frac{1}{1-\beta}})e^{z_\alpha t}, \end{aligned} \tag{6}$$

where $P_0$, $\gamma_0$, $k_0$, and $E(\alpha_0^{\frac{1}{1-\beta}})$ are initial conditions at time 0.

The diffusion path of Internet banking can be derived from Eqs (4) and (6) as

$$F_t = \frac{1}{1 + (\eta y^*_{n,t}/E(y_{n,t}))^{1/g}} = \frac{1}{1 + [\eta y^*_{n,0}/E(y_{n,0})]^{1/g} e^{\frac{1}{g}\{z_k - z_\alpha - z_\gamma - \frac{\beta}{(\beta-1)} z_p\}t}}. \tag{7}$$

We may compare the formula derived in (7) with the classic internal diffusion model (e.g., Griliches 1957, Mansfield 1961), which assumes that the hazard rate of adoption increases with cumulative adoption due to contagion or the "word-of-mouth" effect:

$$\frac{\dot{F}_t}{1 - F_t} = vF_t, \tag{8}$$

where $F_t$ is the fraction of potential adopters who have adopted the innovation at time $t$, and $v$ is a constant contagion parameter. Solving the first-order differential equation (8) yields the logistic function that

$$F_t = \frac{1}{1 + (\frac{1}{F_0} - 1)e^{-vt}}. \tag{9}$$

Variations of this model have been widely used in the economics, sociology, marketing, and management science literature for studying diffusion of innovations (Young, 2009).

Comparing Eq (7) with Eq (9), we find that our formula is indeed equivalent to the logistic diffusion model under very reasonable assumptions. In particular, the model parameters now have clear economic meanings – The so-called "contagion" parameter $v$ is determined by the growth rates of consumer willingness-to-pay, industry deregulation, and technological progress; the initial condition $F_0$ is the fraction of banks that find it profitable to adopt the innovation at the initial time 0:

$$v = (\frac{\beta}{\beta - 1} z_p + z_\gamma + z_\alpha - z_k)/g, \qquad F_0 = \frac{1}{1 + [\eta y^*_{n,0}/E(y_{n,0})]^{1/g}}.$$

Over time, as more banks adopt the innovation, the average bank size keeps rising and the aggregate size distribution of banks shifts towards a new steady state. In the long run, as all banks have adopted the innovation, the cumulative distribution of bank size converges to $G_{y_{a,t}}(x)$ which is again a log-logistic distribution but with a higher mean:

$$G_{y_{a,t}}(x) = 1 - \frac{1}{1 + [\frac{\Gamma(1+g)\Gamma(1-g)}{E(y_{a,t})} x]^{1/g}}, \qquad E(y_{a,t}) = E(y_{n,t})\gamma_t^{\frac{1}{\beta-1}}.$$

Figure 5 illustrates the industry dynamic path. Before Internet banking is introduced, the banking industry stays at a pre-innovation size distribution, drawn with a dotted line. After Internet banking becomes available, in the long run, the banking industry converges to a post-innovation long-run size distribution, drawn with a solid line. In between, the bank size distribution is at a transitional path, drawn with a dashed line. During the transition, at a point in time $t$, there is a size threshold $y^*_{n,t}$, which splits the original size distribution. For banks with size $y_{n,t} \geq y^*_{n,t}$, the size distribution resembles the post-innovation long-run distribution in the range $y_{a,t} \in [\gamma_t^{\frac{1}{\beta-1}} y^*_{n,t}, \infty)$, so $\gamma_t^{\frac{1}{\beta-1}} y^*_{n,t}$ is the minimum size of adopters. Meanwhile, for banks with size $y_{n,t} < y^*_{n,t}$ the size distribution resembles the pre-innovation one, so $y^*_{n,t}$ is the maximum size of non-adopters. Over time, $y^*_{n,t}$ and $\gamma_t^{\frac{1}{\beta-1}} y^*_{n,t}$ fall due to external changes (e.g., demand shift, technological progress, and/or banking deregulation). As a result, Internet banking diffuses into smaller

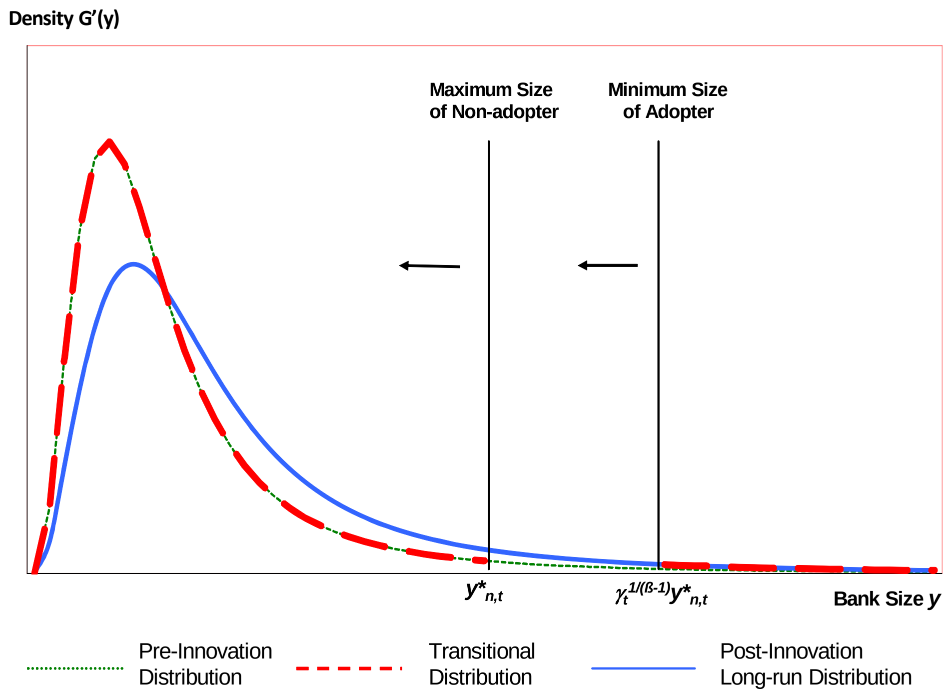


Figure 5: Illustration of the Industry Dynamics

banks, and the bank size distribution gradually converges to the post-innovation long-run distribution.

# 4 Empirical study

In this section, we apply our theory to an empirical study on the diffusion and impact of Internet banking. The sample that we consider includes all in-state banks in each of the 50 U.S. states between 2003-2007. The definitions and summary statistics of our empirical variables are shown in Tables A1 and A2 in the Appendix.

## 4.1 Simultaneous equations

According to our theory, the diffusion and impact of Internet banking can be characterized by two simultaneous equations (an aggregate adoption equation and an average bank size equation) as follows.

Note that the aggregate adoption equation (4) can be rewritten into a log-linear form:

$$g\ln(\frac{F}{1-F}) = -\ln\eta - \ln\frac{\beta}{\beta-1} - \ln k + \ln P + \ln(\gamma^{\frac{1}{\beta-1}} - 1) + \ln E(y_n). \tag{10}$$

An empirical approximation of the average bank size equation (5) can be written as

$$\ln E(y) = \ln E(y_n) + b_1[g\ln(\frac{F}{1-F})] + b_2\ln(\gamma^{\frac{1}{\beta-1}} - 1). \tag{11}$$

Therefore, Eqs (10) and (11) imply

$$g\ln(\frac{F}{1-F}) = a_0 + a_1\ln E(y) + a_1[(1-b_2)\ln(\gamma^{\frac{1}{\beta-1}} - 1) + \ln P - \ln k], \tag{12}$$

where $a_0 = -(\ln\eta + \ln\frac{\beta}{\beta-1})/(1+b_1)$, $a_1 = 1/(1+b_1)$.

Also, Eq (1) suggests

$$y_n = (\frac{P}{\alpha\beta})^{\frac{1}{\beta-1}} \Longrightarrow \ln E(y_n) = \frac{1}{\beta-1}\ln P - \frac{1}{\beta-1}\ln\beta + \ln E(\alpha^{\frac{1}{1-\beta}}).$$

Hence we can rewrite Eq (11) as

$$\ln E(y) = b_0 + b_1[g\ln(\frac{F}{1-F})] + b_2\ln(\gamma^{\frac{1}{\beta-1}} - 1) + \frac{1}{\beta-1}\ln P + \ln E(\alpha^{\frac{1}{1-\beta}}), \tag{13}$$

where $b_0 = \frac{1}{1-\beta}\ln\beta$.

The two equations (12) and (13) are determined simultaneously. Note that the variable $k$ is in Eq (12) but not in (13), and $E(\alpha^{\frac{1}{1-\beta}})$ is in Eq (13) but not in (12). Therefore, they can serve as exclusion restrictions that identify structural parameters.

## 4.2 Empirical specifications

In the empirical study, we estimate the following simultaneous equations based on Eqs (12) and (13) using state-level data of Internet banking diffusion and bank size distribution, where each state is indexed by $j$ and each year is indexed by $t$:[15]

$$g_{j,t}\ln(\frac{F_{j,t}}{1-F_{j,t}}) = a_0 + a_1\ln(E(y)_{j,t}) + \sum_i a_i \ln(X_{i,j,t}) + \sum_l a_l \ln(A_{l,j,t}) + \varepsilon_{j,t}, \quad \text{(Adoption)}$$

$$\ln(E(y)_{j,t}) = b_0 + b_1[g_{j,t}\ln(\frac{F_{j,t}}{1-F_{j,t}})] + \sum_i b_i \ln(X_{i,j,t}) + \sum_l b_l \ln(S_{l,j,t}) + \mu_{j,t}. \quad \text{(Size)}$$

- $F$ is the adoption rate of Internet banking; $g$ is the Gini coefficient of bank size distribution.[16]

- $E(y)$ is the average bank size in terms of deposits.[17]

- $X$ denotes variables shared by both equations, e.g., variables affecting $P$ (price of bank services) and/or $\gamma$ (cost saving due to Internet banking), or variables affecting both $k$ (adoption cost of Internet banking) and $E(\alpha^{\frac{1}{1-\beta}})$ (mean bank productivity).

- $A$ denotes variables only in the *Adoption* equation, e.g., variables affecting $k$ only.

- $S$ denotes variables only in the *Size* equation, e.g., variables affecting $E(\alpha^{\frac{1}{1-\beta}})$ only.

Below is a list of the empirical variables used in our estimation. For most of those variables, we take the log transformation and prefix the variables with "ln" in the notation. Tables A1 and A2 in the Appendix provide more details on each variable.

The dependent variables in the two equations are as follows.

(1) lnTRANODDS_GINI: Log odds ratio for the Internet banking adoption rate adjusted by the Gini coefficient, constructed using the following two variables TRANS – Adoption rate for transactional websites and GINI – Gini coefficient for bank deposits.

[15] Note that our sample includes all in-state banks between 2003-2007 (See footnote 5 for the definition of in-state banks).

[16] Because we do not observe the counterfactual Gini coefficient of bank size distribution in the sample period, we use the sample Gini coefficient as a proxy. Alternatively, we could use the fixed pre-sample Gini coefficient, but the regression results are fairly similar. As shown in Appendix Table A2, the Gini coefficients have large cross-section variation but very small time-series variation.

[17] We also used bank assets as an alternative measure of bank size and the results are very similar.

(2) lnDEPOSITS: Log average bank size, constructed by the variable DEPOSITS – Average bank deposits.

As our theory suggests, we consider three groups of explanatory variables $X$, $A$ and $S$, listed as follows.

$X$ : *Variables in both Adoption and Size equations*

METRO – Ratio of banks in metropolitan areas to all banks.

LOANSPEC – Specialization of lending to consumers.[18]

OFF_DEP – Bank offices per value of deposits.

RMEDFAMINC – Real median family income in 1967 dollars.

POPDEN – Population density.

AGE – Average age of banks.

HHINET – Household Internet access rate.

WAGERATIO – Ratio of computer analyst wage to teller wage.

BHC – Ratio of banks in bank holding companies to total banks.

DEPINT – Ratio of deposits in out-of-state banks to total deposits.

REGION and YEAR – Dummies.

$A$ : *Variables only in Adoption equation*

IMITATE – Years since the first bank in the state adopted a transactional website.

COMRATE – Adoption rate of high-speed Internet among commercial firms in 2003, calculated as an average of urban firms' and rural firms' internet adoption using METRO to weight urban and rural location. Essentially, COMRATE measures in-state banks' exposure to other commercial firms' Internet adoption in each state.

$S$ : *Variables only in Size equation*

DEPOSITS90 – Average bank deposits in 1990.

INTRAREG – A dummy variable for whether the state had intrastate branching restrictions after 1995.

[18] Defined by consumer loans plus 1-4 family mortgages divided by total loans.

Variables in $X$ affect both Internet banking adoption and average bank size. Take HHINET for example: If more households have access to the Internet, local banks may get more cost savings $\gamma$ from adopting Internet banking. However, Internet access also allows households to reach non-local banking services (e.g., out-of-state banks), which may then lower demand and consumer willingness-to-pay $P$ for local banking services. AGE is another example: Established banks typically achieve higher productivity $E(\alpha^{\frac{1}{1-\beta}})$, so they may enjoy an advantage in adopting Internet banking. However, established banks may also face a higher Internet banking adoption cost $k$ compared to young banks given that they have to adapt Internet banking to their legacy computer systems.

The decision on exclusion restrictions $A$ and $S$ is a matter of judgement. We include two variables in $A$: the number of years since the first bank in the state adopted a transactional website (IMITATE) and Internet adoption rate among commercial firms in the state (COMRATE). They are expected to affect the bank size only through their effects on Internet banking adoption. The former variable, IMITATE, is from the *Online Banking Report*, a publication keeping track of the development of Internet banking. The data suggest that the first wave of Internet banking was largely driven by exogenous factors (such as entrepreneurs' risk-taking experiments) rather than cost-benefit calculations assumed in our model. In fact, the correlation between a state's first Internet banking adoption (measured by IMITATE in 2003) and the average bank size in 1990 is -0.001, which suggests IMITATE being a valid instrument. To some extent, this variable may capture the contagion effect suggested by the *internal diffusion* models, but we could also think that a higher value of IMITATE may reduce Internet banking adoption costs by providing more local expertise on bank-specific website design and performance. The latter variable, COMRATE, is constructed based on the information provided by Forman et al (2003). The effect of COMRATE might be ambiguous in theory. On the one hand, a higher value of COMRATE may help Internet banking adoption through the imitation effect. On the other hand, it may delay Internet banking adoption by competing away resources and pushing up local costs of Internet installation and operation. Therefore, we will rely on our empirical estimation to evaluate the overall effect of COMRATE.

We include two variables in $S$: a dummy variable for whether the state had intrastate branching restrictions after 1995 (INTRAREG) and average bank deposits in 1990 (DEPOSITS90). The former value is from Kroszner and Strahan (1999) and the latter is from the Call Report. Both variables are expected to affect the adoption of Internet banking only through their effects on average bank size: INTRAREG may negatively affect the average bank size by imposing high regulation costs; DEPOSITS90 may be positively correlated with current average bank size through the persistence of underlying productivity variables.

## 4.3 Estimation results

Our following discussions focus on the estimation results based on 2SLS (two-stage least squares) models, shown in Tables 1a and 1b. Both the first-stage (reduced-form equation) and the second-stage (structural equation) results are reported. For comparison and robustness checks, we also include in the Appendix the LIML (limited information maximum likelihood) estimation results and the OLS results.

### 4.3.1 Model validation

The 2SLS results suggest that the instrument variables we use are valid and strong. In the first-stage aggregate adoption equation, the coefficients on both lnIMITATE and lnCOMRATE are statistically significant. In the first-stage average bank size equation, the coefficients on INTRAREG and lnDEPOSITS90 have the expected signs and lnDEPOSITS90 is statistically significant.

The relevance of the instruments is also confirmed by F-tests in the first-stage regressions. As a rule of thumb, the F-statistic of a joint test whether all excluded instruments are significant should be bigger than 10 in case of a single endogenous regressor. As shown in Table 1a, this is satisfied in both our adoption and bank size regressions.

Moreover, because we have two instruments for each endogenous variable, we can perform the overidentification test. This test checks whether both instruments are exogenous assuming that at least one of the instruments is exogenous. As shown in Table 1a, the

## Table 1a: Estimated 2SLS Models of Transactional Website Adoption and Size of Bank Deposits

| | Reduced Forms | | Structural Equations | |
|---|---|---|---|---|
| | lnTRANODDS_GINI | lnDEPOSITS | lnTRANODDS_GINI | lnDEPOSITS |
| lnDEPOSITS (fitted) | | | 0.5716<br>(0.0848)*** | |
| lnTRANODDS_GINI (fitted) | | | | 1.3033<br>(0.2686)*** |
| lnIMITATE | 0.3384<br>(0.1506)** | 0.3933<br>(0.2848) | 0.1135<br>(0.1754) | |
| lnCOMRATE | -3.7200<br>(0.7026)*** | -4.9335<br>(1.0055)*** | -0.9002<br>(0.9023) | |
| INTRAREG | -0.0574<br>(0.0493) | -0.1001<br>(0.0764) | | -0.0272<br>(0.0831) |
| lnDEPOSITS90 | 0.2613<br>(0.0463)*** | 0.4572<br>(0.0694)*** | | 0.1164<br>(0.0973) |
| lnMETRO | 0.5357<br>(0.1231)*** | 0.7520<br>(0.2166)*** | 0.1060<br>(0.1636) | 0.0431<br>(0.2536) |
| lnLOANSPEC | 0.1319<br>(0.1191) | 0.3773<br>(0.2138)* | -0.0837<br>(0.1441) | 0.2191<br>(0.1918) |
| lnRMEDFAMINC | -0.3799<br>(0.3451) | 0.2582<br>(0.5425) | -0.5276<br>(0.3653) | 0.7551<br>(0.5659) |
| lnPOPDEN | -0.0490<br>(0.0329) | 0.0994<br>(0.0681) | -0.1059<br>(0.0426)** | 0.1580<br>(0.0616)** |
| lnAGE | -0.2213<br>(0.0872)** | 0.2163<br>(0.1581) | -0.3449<br>(0.1063)*** | 0.4933<br>(0.1668)*** |
| lnHHINET | 2.3160<br>(0.3779)*** | 1.0941<br>(0.6718) | 1.6906<br>(0.3598)*** | -1.9396<br>(0.7602)** |
| lnBHC | 1.2176<br>(0.1804)*** | 1.9964<br>(0.4520)*** | 0.0764<br>(0.2211) | 0.4143<br>(0.4943) |
| lnWGRATIO | -0.3093<br>(0.2177) | -0.5468<br>(0.3983) | 0.0033<br>(0.2575) | -0.1298<br>(0.4067) |
| lnDEPINT | 0.0059<br>(0.0342) | -0.1557<br>(0.0477)*** | 0.0949<br>(0.0327)*** | -0.1626<br>(0.0460)*** |
| lnOFF_DEP | 0.1035<br>(0.0762) | -0.3453<br>(0.1175)*** | 0.3009<br>(0.0851)*** | -0.4823<br>(0.1184)*** |
| Constant | -8.9911<br>(1.3336)*** | -1.2171<br>(2.3079) | -8.2948<br>(1.3169)*** | 10.5330<br>(2.8205)*** |
| Adjusted $R^2$ | 0.83 | 0.78 | 0.75 | 0.74 |
| *N* | 227 | 227 | 227 | 227 |
| Weak instrument test: F(2,201)† | 31.7 | 18.45 | | |
| Exogeneity of regressors-Wald test | | | -4.52*** | -3.24*** |
| Overidentification test: Chi2(1) | | | 0.00 | 0.03 |

* *p*<0.1; ** *p*<0.05; *** *p*<0.01
†Critical values: 19.93 (10%), 11.59 (15%)
Notes: Equations are estimated using two-stage least-squares for the time period 2003 to 2007. Robust standard errors are in parentheses. Estimated coefficients for year and regional dummies are shown in Table 1b.

Table 1b: Estimated 2SLS Models of
Transactional Website Adoption and Size of Bank Deposits
Year and Region Dummy Variables

| | Reduced Forms | | Structural Equations | |
|---|---|---|---|---|
| | lnTRANODDS_GINI | lnDEPOSITS | lnTRANODDS_GINI | lnDEPOSITS |
| d2004 | 0.1068 | -0.0636 | 0.1431 | -0.2087 |
| | (0.0477)** | (0.0975) | (0.0578)** | (0.0908)** |
| d2005 | 0.2408 | -0.0383 | 0.2627 | -0.3630 |
| | (0.0666)*** | (0.1251) | (0.0779)*** | (0.1297)*** |
| d2006 | 0.3517 | -0.1251 | 0.4232 | -0.5983 |
| | (0.0883)*** | (0.1502) | (0.0911)*** | (0.1657)*** |
| d2007 | 0.4693 | -0.1317 | 0.5446 | -0.7626 |
| | (0.1030)*** | (0.1764) | (0.1061)*** | (0.2060)*** |
| Southeast | 0.0849 | 0.2575 | -0.0623 | 0.1442 |
| | (0.0866) | (0.1378)* | (0.1010) | (0.1585) |
| Far west | 0.1203 | 0.9697 | -0.4340 | 0.8207 |
| | (0.0907) | (0.1666)*** | (0.1534)*** | (0.1825)*** |
| Rocky mtn | -0.0450 | 0.3365 | -0.2374 | 0.3965 |
| | (0.0790) | (0.1515)** | (0.0877)*** | (0.1538)*** |
| Southwest | 0.1561 | 0.3933 | -0.0688 | 0.1862 |
| | (0.0942)* | (0.1335)*** | (0.0898) | (0.1537) |
| New England | -0.0632 | 0.3811 | -0.2810 | 0.4719 |
| | (0.1314) | (0.2509) | (0.1406)** | (0.2074)** |
| Mid-east | -0.1308 | -0.3424 | 0.0647 | -0.1675 |
| | (0.1582) | (0.2099) | (0.1527) | (0.2712) |
| Great Lakes | -0.0590 | -0.3125 | 0.1196 | -0.2372 |
| | (0.0700) | (0.1332)** | (0.0871) | (0.1476) |

* $p<0.1$; ** $p<0.05$; *** $p<0.01$

Notes: Equations are estimated using two-stage least-squares for the time period 2003 to 2007. Robust standard errors are in parentheses. Estimated coefficients for other variables in the model equations are in Table 1a.

$\chi^2$ statistics show that we cannot reject the null hypothesis that our instruments are exogenous in either the adoption or the bank size equation.

Finally, we test whether the 2SLS estimates are statistically different from the OLS estimates. The is done by re-running second-stage regressions where the residuals from the first-stage regressions are included (Wooldridge 2010, Chapter 5).[19] This test is robust to heteroscedasticity given that the robust variance estimator is used. The results show that for both the adoption and the bank size equations, the coefficients of the first-stage residuals are statistically significant, which confirm that instrumenting does matter for the estimation.

### 4.3.2 Economic findings

We now turn to the economic findings based on the second-stage estimation results shown in Tables 1a and 1b. Both structural models fit data well, with an $R^2$ of 0.75 for the adoption equation and 0.74 for the bank size equation. Most signs of estimated coefficients, and all of those that are statistically significant, are consistent with our theoretical predictions. The findings are summarized as follows.

In the adoption equation (Table 1a, column 3), the coefficient on the fitted value of lnDEPOSITS is positive and statistically significant. In the size equation (Table 1a, column 4), the coefficient on the fitted value of lnTRANODDS_GINI is also positive and statistically significant. The findings support our theoretical results that Internet banking adoption has a positive causal effect on average bank size, and vice versa. Quantitatively, considering a Gini coefficient equal to 0.57 (the average value in 2003), the results imply that holding everything else constant, a 10 percent increase in average bank size would increase the adoption odds ratio by about 10 percent, and a 10 percent increase of adoption odds ratio would increase the average bank size by about 7.4 percent. To put things into perspective, we may consider a case where the Internet adoption rate is 56.4 percent and the average bank deposits are $311 million, which are mean values of 2003 data. Therefore, based on the 2003 data (Table A2 in the Appendix), a one-

[19] An alternative is to run the Hausman test, but the Hausman test is only valid under homoscedasticity and involves the cumbersome generalized inversion of a non-singular matrix.

standard-deviation increase of average bank deposits from the mean would increase the Internet banking adoption rate from 56.4 percent to 77.1 percent.[20] On the other hand, a one-standard-deviation increase of Internet banking adoption from the mean would raise the average bank deposits from $331 million to $482 million, an increase of 55.0 percent.[21] These findings are in sharp contrast with the OLS regression results (Table A4a in the Appendix, columns 1 and 2). Without addressing the endogeneity of regressors, the OLS results underestimate the impact of lnDEPOSITS and lnTRANODDS_GINI by more than a half.

We also find that Population density (lnPOPDEN) has significant effects on both Internet banking adoption and average bank size. Its effect on Internet banking adoption is negative, suggesting a higher demand for Internet banking in locations with higher cost of travel to bank branches. Its effect on bank size is positive, which confirms that banks in urban areas enjoy more business.

The average bank age in a state (lnAGE) is statistically significant in both equations. The negative coefficient in the adoption equation implies that as the average age of a state's banks increases, the adoption rate falls. This results is consistent with previous findings that *de novo* banks were more likely to adopt Internet banking than incumbent banks (Furst et al. 2001). New banks may find it cheaper to install Internet banking technology in a package with other computer facilities compared to older banks who must add Internet banking to legacy computer systems. Meanwhile, the positive coefficient on lnAGE in the size equation indicates that bank size increases with age, which can be reasonably explained by the accumulation of business expertise and reputation.

Household access to the Internet (lnHHINET) is also statistically significant in both equations. Greater household access to the Internet is associated with a higher adoption of Internet banking, but a smaller average bank size. Both effects are consistent with our discussion above in Section 4.2: If more households have access to the Internet, local banks may get more cost savings from adopting Internet banking. However, Internet access also allows households to reach non-local banking services (e.g., out-of-state banks), so it

[20] This is calculated by solving $F$, where $0.57\times[\ln(\frac{F}{1-F})-\ln(\frac{0.564}{1-0.564})] = 0.5716\times[\ln(311+496)-\ln(311)]$.
[21] This is calculated by solving $y$, where $\ln(y) - \ln(311) = 1.3033 \times 0.57 \times [\ln \frac{0.564+0.136}{1-0.564-0.136} - \ln \frac{0.564}{1-0.564}]$.

negatively affects local bank size.

Competition from out-of-state banks (lnDEPINT) significantly affects Internet banking adoption and in-state bank size. The estimates suggest that more deposits in out-of-state banks push more in-state banks to adopt Internet banking (possibly in order to compete for business). Meanwhile, more competition from out-of-state banks leads to smaller size of in-state banks.

Bank offices per value of deposits (lnOFF_DEP) is statistically significant in both equations. The positive coefficient in the adoption equation implies that banks with more offices may try to explore the synergy between branch banking and Internet banking.[22] The negative coefficient in the size equation suggests that average bank size is smaller where banks have a high number of branches relative to their deposits.

Finally, the year dummies are statistically significant in both equations. After controlling for the other explanatory variables, there is a positive year trend for Internet banking adoption, but a negative year trend for average in-state bank size. In contrast, most regional dummies are not significant or have a negative sign in the adoption equation, in comparison with the excluded PLAIN states which has the lowest Internet banking adoption. This suggests that the observed cross-region differences of Internet adoption are mainly driven by the other explanatory variables in our model rather than the remaining regional fixed effects. We will discuss more on this below.

For robustness checks, we ran a series of additional regressions. First, we used bank assets instead of deposits as an alternative measure of bank size. Second, we explored different samples by looking at state-chartered banks instead of in-state banks or excluding states with a small number of banks (e.g. states with fewer than 10 banks). Third, we employed Fuller's LIML estimators as an alternative way of conducting IV regressions (See Tables A3a and A3b in the Appendix), which have been shown more robust than 2SLS estimators with respect to weak instruments in some recent studies (Murray, 2006). The results are all very similar.[23]

[22] This finding is consistent with optimization of branch network size that compasses both branch-based and non-branch based activities (Hirtle, 2007).

[23] All the robustness check results are available upon request.

### 4.3.3 Regional variations

Our empirical findings above offer useful insights for understanding the diffusion and impact of Internet banking. The results show positive interactions between Internet banking diffusion and average bank size. As explained by our theory, this is because large (more efficient) banks enjoy scale economies of adoption by spreading the fixed adoption cost. Moreover, our findings can help explain the variation in Internet banking diffusion across geographic regions. Particularly, why do the northeast and the west regions have the highest adoption rates, while the central regions have the lowest (See Figure 3)?

Table 2: Mean Values of Selected Variables by Region

(Far West, Plains and New England 2003)

| Variables* | Effect on IB | Far West | Plains | New England |
|---|---|---|---|---|
| OBS (States) | | 6 | 7 | 6 |
| TRANS | | 0.71 | 0.43 | 0.67 |
| GINI | | 0.59 | 0.60 | 0.50 |
| DEPOSITS90 | + | 217.9 | 37.5 | 289.9 |
| IMITATE | + | 5.80 | 6.71 | 6.40 |
| HHINET | + | 61.1 | 55.5 | 60.4 |
| METRO | + | 0.95 | 0.51 | 0.79 |
| BHC | + | 0.66 | 0.87 | 0.62 |
| COMRATE | − | 0.90 | 0.90 | 0.88 |
| AGE | − | 25.6 | 81.6 | 68.1 |

*See Table A1 for variable definitions and sources.

Table 2 presents regional averages of variables that are found significantly affecting Internet banking adoption in the first-stage regression. Far West, Plains and New England are used to represent the west, central and northeast regions respectively.[24] As shown,

[24] Similarly, we can compare the variation in Internet banking diffusion between any other regions. The values of variables for all eight U.S. regions are reported in Table A5 in the Appendix.

the Plains region had a similar Gini coefficient of bank size in 2003 as the Far West and New England, but the Internet banking adoption rate was much lower. Compared with the other two regions, we find that the Plains region has smaller initial bank size, lower household Internet access, fewer banks in metro markets, and older bank vintages. Based on the coefficients (marginal effects) that we uncovered from the first-stage regression, we conclude that these are the factors that have contributed to slow diffusion of Internet banking in the Plains region. On the other hand, our findings reject several alternative hypotheses that may sound appealing, including imitation of early adopters, Internet adoption of commercial firms, and bank holding company membership. In fact, some of those could have been the Plains region's advantage for adoption.

We also rule out several other factors that are only found significantly affecting Internet banking adoption in the second-stage regression, such as deposits held in out-of-state banks, population density, and bank offices per value of deposits. Because those factors show opposite effects on the average bank size in the second-stage regression, their overall effects on Internet banking adoption become insignificant in the first-stage regression where the interaction effects between Internet banking adoption and average bank size are taken into account.

For example, as our second-stage estimation results show, holding everything else constant, an increase of interstate banking competition (measured by lnDEPINT) reduces the average size of in-state banks, but also pushes in-state banks to adopt Internet banking more aggressively. Quantitatively, when we take into account the feedback effects between Internet banking adoption and average bank size, the overall positive effect of lnDEPINT on Internet banking adoption becomes negligible while the overall negative effect on average in-state bank size remains relatively large. To see this more clearly, our second-stage coefficient estimates show that a unit increase of lnDEPINT would directly increase lnTRANODDS_GINI by 0.095 unit, but reduce lnDEPOSITS by 0.163 unit. However, when we take into account the indirect effects through the interactions between lnTRANODDS_ GINI and lnDEPOSITS, the final effect on lnTRANODDS_GINI is reduced to less than 0.01 unit, and the final effect on lnDEPOSITS remains more than

0.15 unit.[25] This is consistent with the coefficient estimates obtained in our first-stage regressions. Similarly, population density (POPDEN) and bank offices per value of deposits (OFF_DEP) each affect in-state banks in terms of their size and Internet banking adoption but the effects offset one another. These variables thus become unimportant in explaining the regional variation in Internet banking diffusion.

## 4.4 Internal *vs.* external diffusion

Our empirical analysis sheds light on the debate regarding internal and external diffusion models. The internal diffusion models predict $S$-shaped logistic diffusion curves, which could serve as a convenient tool for data fitting or forecasting. In fact, one could use our state-level Internet banking diffusion data to estimate Eq (9) or its simple log-linear transformation:

$$\ln\left(\frac{F_{j,t}}{1-F_{j,t}}\right) = a_j + v_j t. \tag{14}$$

Accordingly, the diffusion process in a state $j$ can be explained by the estimates of two state-specific parameters: the initial condition $a_j$ and the contagion rate $v_j$. However, it is difficult to explore deeper economic questions beyond that, for example, why the contagion rate, or the "word-of-mouth" effect, differs across regions, and why large banks rather than small banks tend to be the early adopters.

In contrast, the external-diffusion approach we take in this paper provides a better micro-founded explanation. By modelling explicitly the size heterogeneity of banks, we keep the appealing feature of $S$-shaped logistic diffusion curves but connect them to more meaningful economic factors. Our empirical findings, besides providing good fitting of the data, offer several important implications:

- First, employing instrument variables in the estimation confirms the causal effect of firm size distribution on technology diffusion, which justifies the external-diffusion approach we take.

[25] Using the second-stage coefficient estimates, we can solve the simultaneous equations and get the overall effects of lnDEPINT: $\partial$(lnTRANODDS_GINI)/$\partial$(lnDEPINT)$=\frac{-0.1626\times0.5716+0.0949}{1-0.5716\times1.3033} = 0.0077$; while $\partial$(lnDEPOSITS)/$\partial$(lnDEPINT) $= \frac{1.3033\times0.0949-0.1626}{1-0.5716\times1.3033} = -0.1526$.

- Second, the variation in diffusion rates across regions can be well explained by underlying technological, economic, and institutional factors. We find that, after controlling for those variables in the regressions, regional dummies are left with little explanatory power.

- Finally, technology diffusion and firm size distribution are jointly determined, so they should not be treated exogenously to each other. As our results show, without addressing the endogeneity problem, the OLS regression results can be much biased.

# 5 Conclusion

Taking Internet banking as an example, we study diffusion of cost-saving technological innovations and the impact on firm size distribution. Our theory suggests that when such an innovation is initially introduced, large firms enjoy cost advantages in adopting it early on and thus increasing their size relative to non-adopters. Over time, due to external changes (e.g., demand shift, technological progress, and/or industry deregulation), the innovation gradually diffuses into smaller firms. As a result, the firm size distribution shifts towards a new steady state, and the technology diffusion follows an $S$-shaped logistic curve. During the process, there exist important interaction effects between technology diffusion and firm size distribution.

Applying the theory to an empirical study of Internet banking diffusion among banks across 50 U.S. states, we examine the technological, economic and institutional factors governing the process. The empirical findings allow us to disentangle the interrelationship between Internet banking diffusion and bank size distribution, and explain the variation in diffusion rates across geographic regions.

Our analysis bridges a gap between internal and external diffusion models. The approach that we develop goes beyond the Internet banking application. It provides a natural framework for studying the interaction effects between technology diffusion and firm size distribution, which can potentially be applied to other episodes of technology diffusion.

## References


[1] Bass, F. M. (1969). "A New Product Growth for Model Consumer Durables," *Management Science*,15(5), 215-227.

[2] Bass, F. M. (2004). "Comments on "A New Product Growth for Model Consumer Durables": The Bass Model," *Management Science* 50(12), 1833-1840.

[3] Berger, Allen, Anil Kashyap and Joseph Scalise (1995). "The Transformation of the U.S. Banking Industry: What a Long, Strange Trip It's Been." *Brookings Papers on Economic Activity* 2: 55–201.

[4] Cabral, Luis, and Jose Mata (2003). "On the Evolution of the Firm Size Distribution: Facts and Theory," *American Economic Review*, 1075-1090.

[5] Courchane, Marsha, David Nickerson and Richard J. Sullivan (2002). "Investment in Internet Banking as a Real Option: Theory and Tests," *The Journal of Multinational Financial Management*, 12(4-5), 347-63.

[6] David, Paul A. (1969). "A Contribution to the Theory of Diffusion," Stanford University Research Center in Economic Growth Memorandum 71.

[7] David, Paul A. (2005). "Zvi Griliches on Diffusion, Lags, and Productivity Growth…Connecting the Dots." Labor and Demography Working Paper 0502012.

[8] DeYoung, R., W.W. Lang, D.L. Nolle (2007). "How the Internet Affects Output and Performance at Community Banks," *Journal of Banking and Finance,* 31, 1033-1060.

[9] Ennis, Huberto (2001). "On the Size Distribution of Banks," *Federal Reserve Bank of Richmond Economic Quarterly,* 87: 1-25.

[10] Forman, C., A. Goldfarb and S. Greenstein (2003). "The Geographic Dispersion of Commercial Internet Use," *Rethinking Rights and Regulations: Institutional Responses to New Communication Technologies*, Cambridge: MIT Press, 113-45.

[11] Furst, Karen, William Lang and Daniel Nolle (2001). "Internet Banking in the US:

Landscape, Prospects, and Industry Implications," *Journal of Financial Transformation*, 2, 45-52.

[12] Griliches, Zvi (1957). "Hybrid Corn: An Exploration in the Economics of Technological Change," *Econometrica*, 25, 501-522.

[13] Hernández-Murillo, Rubén, Gerard Llobet, Roberto Fuentes (2010). "Strategic Online Banking Adoption," *Journal of Banking and Finance*, 34(7), 1650-1663.

[14] Hirtle, Beverly (2007). "The Impact of Network Size on Bank Branch Performance," *Journal of Banking and Finance*, 31, 3782-3805.

[15] Hopenhayn, H. A. (1992). "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 60(2): 1127-1150.

[16] Janicki, H. and E. S. Prescott (2006). "Changes in the Size Distribution of U.S. Banks: 1960-2005," *Federal Reserve Bank of Richmond Economic Quarterly*, 92, 291-316.

[17] Jovanovic, Boyan (1982). "Selection and Evolution of Industry," *Econometrica*, 50(3), 649-670.

[18] Jones, K. and T. Critchfield (2005). "Consolidation in the U.S. Banking Industry: Is the 'Long, Strange Trip'About to End?" *FDIC Banking Review* 17: 31-61.

[19] Kroszner, Randall and Philip Strahan (1999). "What Drives Deregulation? Economics and Politics of the Relaxation of Banking Branching Restrictions," *Quarterly Journal of Economics*, 114(4), 1437-1467.

[20] Lucas, Robert (1978). "On the Size Distribution of Business Firms," *Bell Journal of Economics*, 9(2), 508-23.

[21] Mansfield, Edwin (1961). "Technical Change and the Rate of Innovation," *Econometrica*, 29, 741-66.

[22] McCord, R. and E. S. Prescott (2014). "The Financial Crisis, the Collapse of Bank Entry, and Changes in the Size Distribution of Banks," *Federal Reserve Bank of Richmond Economic Quarterly*, 100, 23-50.

[23] Murray, Michael (2006). "Avoiding Invalid Instruments and Coping with Weak Instruments," *Journal of Economic Perspective*, 20(4), 111-132.

[24] Rogers, Everett M. (2003). *Diffusion of Innovations*. 5th ed. New York: Free Press.

[25] Stoneman, Paul (2002). *The Economics of Technological Diffusion*. Oxford: Blackwell.

[26] Sutton, J. (1997). "Gibrat's Legacy," *Journal of Economic Literature*, 35(1), 40-59.

[27] Wang, Zhu (2007). "Technological Innovation and Market Turbulence: The Dot-com Experience," *Review of Economic Dynamics*, 10(1), 78-105.

[28] Wooldridge, Jeffrey M. (2010). *Econometric Analysis of Cross Section and Panel Data*, MIT Press.

[29] Young H. P. (2009). "Innovation Diffusion in Heterogeneous Populations: Contagion, Social influence, and Social learning." *American Economic Review*, 99, 1899-1924.

Table A1: Empirical Variable Definitions and Sources

| Variable name | Definition | Source |
|---|---|---|
| TRANS | Adoption rate for transactional websites | Call Report |
| TRANODDS | Odds ratio for adoption of transactional websites | Call Report |
| GINI | Gini coefficient for bank deposits | Call Report |
| DEPOSITS | Average bank deposits | Call Report |
| METRO | Ratio of banks in metropolitan areas to all banks | Call Report |
| LOANSPEC | Specialization of lending to consumers (consumer loans plus 1-4 family mortgages / total loans) | Call Report |
| OFF_DEP | Bank offices per value of deposits | Call Report; FDIC *Summary of Deposits* |
| RMEDFAMINC | Median family income (in 1967 dollars) | U.S. Census Bureau |
| POPDEN | Population density | *Statistical Abstract of the United States* |
| IMITATE | Years since the first bank in the state adopted a transactional website | *Online Banking Report* |
| AGE | Average age of banks | Call Report |
| HHINET | Household access rate for Internet | *Statistical Abstract of the United States* |
| WGRATIO | Ratio of computer analyst wage to teller wage | Bureau of Labor Statistics |
| INTRAREG | Indicator variable for whether the state had branching restrictions after 1995 | Krozner and Strahan, 1999 |
| BHC | Ratio of banks in bank holding companies to total banks | Call Report |
| DEPINT | Ratio of deposits in out-of-state banks to total deposits | FDIC *Summary of Deposits* |
| COMRATE | Adoption rate of high-speed internet among commercial firms | Forman, et.al., 2003 |
| DEPOSITS90 | Average bank deposits in 1990 | Call Report |
| Regional dummy variables: | | Bureau of Economic Analysis |
| SE | Southeast: AL, AR, FL, GA, KY, LA, MS, NC, SC, TN, VA, WV | |
| FARWEST | Far West: AK, CA, HI, NV, OR, WA | |
| ROCKYMTN | Rocky Mountain: CO, ID, MT, UT, WY | |
| PLAINS | Plains: IA, KS, MN, MO, NE ,ND, SD | |
| SW | Southwest: AZ, NM, OK, TX | |
| NWENGLND | New England: CT, MA, ME, NH, RI, VT | |
| MIDEAST | Middle East: DC, DE, MD, NJ, NY, PA | |
| GRTLAKE | Great Lakes: IL, IN, MI, OH, WI | |

Notes: Data are for individual states.

Variables for banks are unweighted averages for those located in individual states. Selected banks are full-service, retail commercial banks.

Data for adoption of high-speed internet among commercial firms is for 2003. COMRATE is an average of urban firms' and rural firms' internet adoption, using METRO to weight urban and rural location.

BEA Regions are a set of Geographic Areas that are aggregations of the states. The regional classifications, which were developed in the mid-1950s, are based on the homogeneity of the states in terms of economic characteristics, such as the industrial composition of the labor force, and in terms of demographic, social, and cultural characteristics. For a brief description of the regional classification of states used by BEA, see U.S. Department of Commerce, Census Bureau, Geographic Areas Reference Manual, Washington, DC, U.S. Government Printing Office, November 1994, pp. 6-18;6-19.

Table A2: Summary Statistics

| | 2003 | | | | 2005 | | | | 2007 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLE | Mean | S. D. | Min | Max | Mean | S. D. | Min | Max | Mean | S. D. | Min | Max |
| TRANS | 0.564 | 0.136 | 0.263 | 0.852 | 0.729 | 0.121 | 0.456 | 0.949 | 0.830 | 0.095 | 0.624 | 0.978 |
| TRANODDS | 1.588 | 1.063 | 0.357 | 5.750 | 3.880 | 3.474 | 0.837 | 18.501 | 7.888 | 7.930 | 1.659 | 44.004 |
| GINI | 0.574 | 0.122 | 0.338 | 0.847 | 0.568 | 0.119 | 0.325 | 0.862 | 0.583 | 0.117 | 0.305 | 0.908 |
| DEPOSITS* | $311 | $496 | $65 | $3,307 | $406 | $783 | $67 | $4,028 | $486 | $970 | $71 | $5,057 |
| METRO | 0.741 | 0.187 | 0.295 | 1.000 | 0.741 | 0.184 | 0.300 | 1.000 | 0.737 | 0.185 | 0.298 | 1.000 |
| LOANSPEC | 0.373 | 0.121 | 0.144 | 0.608 | 0.351 | 0.122 | 0.137 | 0.581 | 0.333 | 0.124 | 0.102 | 0.574 |
| OFF_DEP | 0.023 | 0.008 | 0.003 | 0.037 | 0.021 | 0.008 | 0.003 | 0.034 | 0.020 | 0.008 | 0.003 | 0.031 |
| RMEDFAMINC** | $93.5 | $13.9 | $70.0 | $126.9 | $93.6 | $13.9 | $70.0 | $129.2 | $95.3 | $12.6 | $72.1 | $131.1 |
| POPDEN | 148.0 | 179.6 | 1.1 | 821.4 | 153.4 | 181.5 | 5.2 | 820.6 | 148.9 | 175.8 | 5.4 | 822.7 |
| IMITATE | 6.745 | 1.132 | 4.000 | 9.000 | 8.783 | 1.114 | 6.000 | 11.000 | 10.791 | 1.103 | 8.000 | 13.000 |
| AGE | 58.7 | 23.2 | 6.7 | 111.7 | 59.2 | 23.9 | 7.4 | 112.5 | 60.4 | 25.6 | 5.8 | 121.5 |
| HHINET | 54.4 | 6.2 | 38.9 | 67.6 | 57.6 | 6.1 | 42.4 | 70.1 | 60.7 | 6.2 | 46.0 | 71.6 |
| WGRATIO | 3.035 | 0.238 | 2.417 | 3.396 | 3.056 | 0.218 | 2.689 | 3.497 | 3.049 | 0.268 | 2.230 | 3.572 |
| INTRAREG | 0.234 | 0.428 | 0.000 | 1.000 | 0.239 | 0.431 | 0.000 | 1.000 | 0.256 | 0.441 | 0.000 | 1.000 |
| BHC | 0.776 | 0.118 | 0.444 | 0.931 | 0.792 | 0.121 | 0.429 | 0.937 | 0.808 | 0.110 | 0.579 | 0.940 |
| DEPINT | 0.283 | 0.185 | 0.002 | 0.741 | 0.351 | 0.197 | 0.005 | 0.843 | 0.341 | 0.192 | 0.020 | 0.831 |
| COMRATE | 0.889 | 0.026 | 0.778 | 0.921 | 0.889 | 0.026 | 0.777 | 0.922 | 0.889 | 0.027 | 0.776 | 0.922 |
| DEPOSITS90* | $207 | $365 | $26 | $2.393 | $207 | $369 | $26 | $2,393 | $209 | $382 | $26 | $2,393 |
| SE | 0.255 | 0.441 | 0 | 1 | 0.261 | 0.444 | 0 | 1 | 0.279 | 0.454 | 0 | 1 |
| FARWEST | 0.106 | 0.312 | 0 | 1 | 0.087 | 0.285 | 0 | 1 | 0.093 | 0.294 | 0 | 1 |
| ROCKYMTN | 0.106 | 0.312 | 0 | 1 | 0.109 | 0.315 | 0 | 1 | 0.093 | 0.294 | 0 | 1 |
| SW | 0.085 | 0.282 | 0 | 1 | 0.087 | 0.285 | 0 | 1 | 0.093 | 0.294 | 0 | 1 |
| NWENGLND | 0.106 | 0.312 | 0 | 1 | 0.109 | 0.315 | 0 | 1 | 0.093 | 0.294 | 0 | 1 |
| MIDEAST | 0.085 | 0.282 | 0 | 1 | 0.087 | 0.285 | 0 | 1 | 0.070 | 0.258 | 0 | 1 |
| GRTLAKE | 0.106 | 0.312 | 0 | 1 | 0.109 | 0.315 | 0 | 1 | 0.116 | 0.324 | 0 | 1 |
| PLAINS | 0.149 | 0.360 | 0 | 1 | 0.152 | 0.363 | 0 | 1 | 0.163 | 0.374 | 0 | 1 |

Notes: Sample population includes the 50 states in the U.S. and the District of Columbia. The sample size varies from year to year because the transactional website adoption rate reached 100% for some observations and TRANODDS cannot be calculated. The actual sample size in 2003, 2005, and 2007 is 47, 46, and 43.
See Table A1 for variable definitions and sources.
*In millions.
**In thousands

Table A3a: Estimated LIML Models of
Transactional Website Adoption and Size of Bank Deposits

| | Structural Equations | |
|---|---|---|
| | lnTRANODDS_GINI | lnDEPOSITS |
| lnDEPOSITS (fitted) | 0.5716<br>(0.0820)*** | |
| lnTRANODDS_GINI (fitted) | | 1.3040<br>(0.3192)*** |
| lnIMITATE | 0.1135<br>(0.1872) | |
| lnCOMRATE | -0.9002<br>(0.8624) | |
| INTRAREG | | -0.0272<br>(0.0994) |
| lnDEPOSITS90 | | 0.1162<br>(0.0956) |
| lnMETRO | 0.1060<br>(0.1619) | 0.0428<br>(0.2556) |
| lnLOANSPEC | -0.0837<br>(0.1313) | 0.2190<br>(0.1856) |
| lnRMEDFAMINC | -0.5276<br>(0.2979)* | 0.7553<br>(0.4753) |
| lnPOPDEN | -0.1059<br>(0.0398)*** | 0.1581<br>(0.0578)*** |
| lnAGE | -0.3449<br>(0.0779)*** | 0.4935<br>(0.1414)*** |
| lnHHINET | 1.6906<br>(0.3632)*** | -1.9412<br>(0.8974)** |
| lnBHC | 0.0764<br>(0.2032) | 0.4135<br>(0.4498) |
| lnWGRATIO | 0.0033<br>(0.2708) | -0.1295<br>(0.4423) |
| lnDEPINT | 0.0949<br>(0.0293)*** | -0.1626<br>(0.0418)*** |
| lnOFF_DEP | 0.3009<br>(0.0743)*** | -0.4824<br>(0.1017)*** |
| Constant | -8.2948<br>(1.3989)*** | 10.5383<br>(3.1675)*** |
| Adjusted $R^2$ | 0.75 | 0.74 |
| *N* | 227 | 227 |
| Weak instrument test: F(2, 201) † | 31.7 | 15.9 |

* $p$<0.1; ** $p$<0.05; *** $p$<0.01

†Critical values: 8.68 (10%), 5.33 (15%).

Notes: Equations are estimated using limited information maximum likelihood for the time period 2003 to 2007. Estimated coefficients for year and regional dummy variables are shown in Table A3b.

## Table A3b: Estimated LIML Models of Transactional Website Adoption and Size of Bank Deposits Year and Regional Dummy Variables

| | Structural Equations | |
|---|---|---|
| | lnTRANODDS_GINI | lnDEPOSITS |
| d2004 | 0.1431 | -0.2088 |
| | (0.0600)** | (0.1004)** |
| d2005 | 0.2627 | -0.3633 |
| | (0.0731)*** | (0.1414)** |
| d2006 | 0.4232 | -0.5987 |
| | (0.0866)*** | (0.1855)*** |
| d2007 | 0.5446 | -0.7631 |
| | (0.1032)*** | (0.2333)*** |
| Southeast | -0.0623 | 0.1440 |
| | (0.1104) | (0.1744) |
| Far west | -0.4340 | 0.8206 |
| | (0.1444)*** | (0.1724)*** |
| Rocky mtn | -0.2374 | 0.3965 |
| | (0.0948)** | (0.1473)*** |
| Southwest | -0.0688 | 0.1861 |
| | (0.1152) | (0.1889) |
| New England | -0.2810 | 0.4718 |
| | (0.1539)* | (0.2050)** |
| Mid-east | 0.0647 | -0.1677 |
| | (0.1422) | (0.2312) |
| Great Lakes | 0.1196 | -0.2373 |
| | (0.1006) | (0.1614) |

* $p$<0.1; ** $p$<0.05; *** $p$<0.01

Notes: Equations are estimated using limited information maximum likelihood for the time period 2003 to 2007. Estimated coefficients for other variables in the model equations are in Table A3a.

## Table A4a: Estimated OLS Models of Transactional Website Adoption and Size of Bank Deposits

| | Structural Equations | |
|---|---|---|
| | lnTRANODDS_GINI | lnDEPOSITS |
| lnDEPOSITS | 0.2467<br>(0.0436)*** | |
| lnTRANODDS_GINI | | 0.5674<br>(0.1390)*** |
| lnIMITATE | 0.1915<br>(0.1530) | |
| lnCOMRATE | -2.0247<br>(0.7779)*** | |
| INTRAREG | | -0.0628<br>(0.0743) |
| lnDEPOSITS90 | | 0.2873<br>(0.0734)*** |
| lnMETRO | 0.3926<br>(0.1280)*** | 0.2939<br>(0.2201) |
| lnLOANSPEC | 0.0511<br>(0.1205) | 0.3086<br>(0.1851)* |
| lnRMEDFAMINC | -0.4229<br>(0.3247) | 0.5603<br>(0.5378) |
| lnPOPDEN | -0.0844<br>(0.0324)*** | 0.1115<br>(0.0544)** |
| lnAGE | -0.3696<br>(0.0928)*** | 0.2814<br>(0.1488)* |
| lnHHINET | 1.7774<br>(0.3507)*** | -0.3372<br>(0.6960) |
| lnBHC | 0.5697<br>(0.1616)*** | 1.2538<br>(0.4409)*** |
| lnWGRATIO | -0.1073<br>(0.2257) | -0.5205<br>(0.3756) |
| lnDEPINT | 0.0487<br>(0.0281)* | -0.1614<br>(0.0434)*** |
| lnOFF_DEP | 0.1244<br>(0.0629)** | -0.4006<br>(0.1104)*** |
| Constant | -5.8434<br>(1.1062)*** | 5.2588<br>(2.3939)** |
| Adjusted $R^2$ | 0.82 | 0.79 |
| *N* | 227 | 227 |

* $p$<0.1; ** $p$<0.05; *** $p$<0.01

Notes: Equations are estimated using ordinary least squares for the time period 2003 to 2007. Robust standard errors are in parentheses. Estimated coefficients for year and regional dummy variables are shown in Table A4b.

Table A4b: Estimated OLS Models of
Transactional Website Adoption and Size of Bank Deposits
Year and Region Dummy Variables

| | Structural Equations | |
|---|---|---|
| | lnTRANODDS_GINI | lnDEPOSITS |
| d2004 | 0.1362 | -0.0824 |
| | (0.0482)*** | (0.0866) |
| d2005 | 0.2750 | -0.0949 |
| | (0.0658)*** | (0.1071) |
| d2006 | 0.4246 | -0.2122 |
| | (0.0820)*** | (0.1273)* |
| d2007 | 0.5507 | -0.2509 |
| | (0.0980)*** | (0.1349)* |
| Southeast | 0.0847 | 0.2411 |
| | (0.0850) | (0.1412)* |
| Far west | -0.0500 | 0.8825 |
| | (0.1094) | (0.1566)*** |
| Rocky mtn | -0.1454 | 0.3632 |
| | (0.0712)** | (0.1457)** |
| Southwest | 0.0829 | 0.3457 |
| | (0.0796) | (0.1326)*** |
| New England | 0.0842 | 0.3914 |
| | (0.1112) | (0.2245)* |
| Mid-east | 0.2995 | -0.2680 |
| | (0.1223)** | (0.2222) |
| Great Lakes | 0.1716 | -0.2620 |
| | (0.0731)** | (0.1356)* |

* $p<0.1$; ** $p<0.05$; *** $p<0.01$

Notes: Equations are estimated using ordinary least-squares for the time period 2003 to 2007. Robust standard errors are in parentheses. Estimated coefficients for other variables in the model equation are in Table A4a.

Table A5: Mean Values of Selected Variables by Region 2003

| VARIABLE | New England | Mideast | Southeast | Great Lakes | Plains | Rocky Mountain | Southwest | Far West |
|---|---|---|---|---|---|---|---|---|
| TRANS | 0.666 | 0.689 | 0.525 | 0.534 | 0.427 | 0.561 | 0.532 | 0.706 |
| TRANODDS | 2.031 | 2.476 | 1.267 | 1.225 | 0.801 | 1.382 | 1.390 | 3.031 |
| GINI | 0.495 | 0.668 | 0.514 | 0.668 | 0.596 | 0.485 | 0.699 | 0.585 |
| DEPOSITS* | 429.7 | 1152.9 | 190.7 | 257.1 | 101.3 | 131.6 | 251.5 | 378.2 |
| METRO | 0.794 | 0.936 | 0.708 | 0.769 | 0.509 | 0.681 | 0.766 | 0.949 |
| LOANSPEC | 0.475 | 0.481 | 0.441 | 0.459 | 0.294 | 0.279 | 0.320 | 0.179 |
| OFF_DEP | 0.019 | 0.014 | 0.026 | 0.021 | 0.028 | 0.025 | 0.023 | 0.019 |
| RMEDFAMINC** | 109.7 | 107.5 | 82.2 | 97.9 | 93.2 | 92.4 | 81.5 | 100.3 |
| POPDEN | 358.6 | 416.2 | 132.3 | 191.5 | 39.2 | 20.1 | 50.0 | 75.6 |
| IMITATE | 6.400 | 7.500 | 7.000 | 7.800 | 6.714 | 6.000 | 6.500 | 5.800 |
| AGE | 68.1 | 64.8 | 53.6 | 78.6 | 81.6 | 47.9 | 46.3 | 25.6 |
| HHINET | 60.4 | 56.0 | 48.6 | 52.8 | 55.5 | 58.0 | 50.0 | 61.1 |
| WGRATIO | 2.884 | 3.209 | 3.015 | 3.183 | 3.125 | 2.905 | 3.074 | 2.922 |
| INTRAREG | 0.000 | 0.000 | 0.250 | 0.000 | 0.571 | 0.600 | 0.250 | 0.000 |
| BHC | 0.621 | 0.768 | 0.785 | 0.854 | 0.873 | 0.822 | 0.774 | 0.656 |
| DEPINT | 0.324 | 0.224 | 0.313 | 0.184 | 0.164 | 0.305 | 0.379 | 0.382 |
| COMRATE | 0.883 | 0.880 | 0.889 | 0.902 | 0.898 | 0.866 | 0.885 | 0.901 |
| DEPOSITS90* | 289.9 | 985.5 | 116.7 | 118.0 | 37.5 | 63.7 | 169.6 | 217.9 |
| | | | | | | | | |
| OBS (States) | 6 | 5 | 12 | 5 | 7 | 5 | 4 | 6 |

Notes: See Table A1 for variable definitions and sources. See Table A2 for the national average of variables.
*In millions. **In thousands