Quantifying Equilibrium Network Externalities in the ACH Banking Industry[1]

Daniel A. Ackerberg
Department of Economics
University of Arizona
and NBER
daniel_ackerberg@nber.org

and

Gautam Gowrisankaran
John M. Olin School of Business
Washington University in St. Louis
and NBER
gautam_gowrisankaran@nber.org

This Version: March 1, 2004

Abstract

We seek to estimate the causes and magnitudes of network externalities for the automated clearinghouse (ACH) electronic payments system, using a panel data set on individual bank usage of ACH. We construct an equilibrium model of consumer and bank adoption of ACH in the presence of a network. The model identifies network externalities from correlations of changes in usage levels for banks within a network, from changes in usage following changes in market concentration or sizes of competitors and from the adoption decisions of banks outside the network with small branches in the network. The model can separately identify consumer and bank network effects. Using a dataset of localized networks, we structurally estimate the parameters of the model by matching equilibrium behavior to the data using an indirect inference method of simulated moments procedure. The parameters are estimated with good precision and fit various moments of the data reasonably well. We find that most of the impediment to ACH adoption is due to large consumer fixed costs of adoption. Policies to provide moderate subsidies to consumers and larger subsidies to banks for ACH adoption would increase consumer and bank welfare significantly.

[1] We acknowledge funding from the NET Institute and the National Science Foundation (Grant SES-0318170), and thank Steve Berry, Jinyong Hahn, Brian McManus, Andrea Moro, Klaas van't Veld and seminar participants at numerous institutions for helpful comments.
.

1. Introduction

The goal of this paper is to estimate the size and importance of network externalities for the automated clearinghouse (ACH) banking industry using an equilibrium model of ACH usage. ACH is an electronic payment mechanism developed by the Federal Reserve and used by banks. ACH is a network and a two-sided market: banks on both sides of a transaction must adopt ACH technology for an ACH transaction to occur. Network externalities are thought to exist in many high-technology industries. Examples include fax machines, where network effects may exist because two separate parties must communicate for a transaction to occur and computers, where network effects may exist because information on how to use new technology is costly. ACH shares the network features of fax machines, computers and other technological goods, and hence network externalities may exist for ACH.

If present, network externalities typically cause underutilization of the network good. Network externalities can also generate multiple equilibria. When the network externality is positive, these Nash equilibria can be Pareto ranked, and it is possible that the industry is stuck in a Pareto inferior equilibrium, characterized by even less usage than the Pareto best equilibrium. The underutilization is particularly relevant for the case of ACH. In an age when computers and technology have become prevalent, most payments continue to be performed with checks and cash. By estimating the magnitude of network externalities, we can further understand the causes of such externalities, uncover how much the usage of ACH differs from the socially optimal level, and find out whether markets are stuck in Pareto inferior equilibria. Moreover, by estimating an equilibrium model, we can evaluate the welfare and usage consequences of policies such as government subsidies of the network good.

This work extends previous research on estimating network externalities for ACH (Gowrisankaran and Stavins, 2004). That study postulated a simple game which resulted in bank adoption of ACH being a function of the adoption decisions of other banks in the network, as well as of a bank's characteristics, such as its size. The interdependence of preferences for ACH adoption leads to a simultaneity in the equilibrium adoption decisions of banks, making identification of the network externalities potentially difficult. Thus, the study proposed three methods to identify network externalities: examining whether adoption is clustered (after controlling for bank fixed effects), using excluded exogenous variables based on bank size to control for endogenous adoption decisions, and exploiting the quasi-experimental variation from the adoption decisions of small, remote branches of banks. Each of the three strategies revealed significant and positive network externalities, even after controlling for factors such as economies of scale and market power.

This paper builds on the previous research, by specifying and structurally estimating an equilibrium model of technology adoption for ACH in the presence of network externalities. The estimation uses similar data to the earlier work, and hence is identified from the same sources. However, our use of structural estimation has several advantages. First, we estimate a functional form for the network model that is directly consistent with the underlying theory of consumer utility maximization. Most importantly, this allows us to identify whether the network effects are arising at the consumer or bank level. Additionally, this allows us to efficiently combine data on bank adoption of ACH and volume conditional on adoption and to handle networks with one bank in a logical way.[2] Second, we can recover the magnitudes of the network externalities, in a way that uses the power from the combination of all three methods of identification.[3] Third, the structural model leads very naturally to welfare and policy analysis. Note that the empirical distinction between consumer and bank level network externalities is very important here. With a subsidy to promote adoption, for example, one would want to know whom to subsidize, banks or

---

[2] In contrast, the earlier work could either model adoption or bank volume as dependent variables. There were measurement issues in using the adoption variable, but it is difficult to model the quantity choice outside of a structural model. Moreover, the previous work had to exclude networks with one bank, because the network variables, which are based on the fraction of other banks adopting ACH, were not defined for this case.

[3] The earlier work was more successful at hypothesis testing *whether* network externalities existed rather than actually measuring the magnitude of them. It was able to recover magnitudes of the network externalities for a few of the individual specifications. However, these specifications were somewhat limited and problematic. For instance, the quasi-experimental source of variation identified the magnitudes of network externalities, but only for a very small data set (0.2% of the total observations) of rural banks. The instrumental variables specification identified the network externalities but at the cost of imposing a linear functional form for the discrete adoption variable. The work on treatment effects (e.g. Heckman and Robb (1987) and Angrist and Imbens (1994)) suggests that linear probability models cause significant problems identifying causal effects using IV because of heteroskedasticity. The correlation source of variation could not be used at all to identify the magnitude of the network externalities without

consumers. Last, the structural estimation methods that we develop here are novel and contribute to the literature on structural estimation of simultaneous games and network games in particular.[4]

We specify a two-sided market model of technology adoption as follows. We consider a localized, repeated static market with a given set of banks and consumers that are tied to these banks. Each consumer must make a fixed number of transactions to other consumers evenly distributed throughout the network; these transactions can be made using either checks or ACH. While all banks and consumers accept checks, some may not have adopted ACH. Some banks are local to the market while others are branches of big banks based outside the network. In each time period, local banks decide whether to adopt ACH capabilities, based on whether the marginal profits from ACH transactions conditional on adoption are greater than the fixed costs of adoption; the decisions of non-local banks are made exogenously and known to the local banks.

Following bank adoption, each consumer at each bank that has adopted ACH chooses whether or not to adopt ACH. If the consumer adopts ACH, she must pay a fixed cost of adoption, but then can, and by assumption will, use ACH for her transactions to those consumers that have also adopted ACH. Some ACH transactions only occur if *both* consumers have adopted ACH; we call these "two-way" transactions. These would include business-to-business transactions, for instance. We also allow some ACH transactions to occur when the recipient consumer has *not* formally adopted ACH technology. These are called "one-way" transactions, and correspond to transactions such as direct deposit paychecks and online mortgage payments where the employee or homeowner has not formally adopted ACH. We specify and estimate the proportion of one-way transactions as a structural parameter of the model. We model the consumer fixed costs of adoption with random effects to control for correlated preferences. The fact that two-way ACH transactions can only be made to other individuals who have adopted

structural methods. All of the identification of the magnitudes used a reduced-form profit function for banks that was not consistent with the underlying consumer preferences.

[4] For instance, Brock and Durlauf (2001) discusses identification for social interaction games, which are conceptually identical to network externality games. Topa (2001) structurally estimates a social interaction model using a GMM procedure. We develop an indirect inference estimation procedure for our model, that can be used to estimate these types of games.

implies that, in equilibrium, consumers are more likely to adopt if more consumers have adopted ACH. Similarly, banks are more likely to adopt ACH if more consumers and banks are expected to adopt. There may be multiple equilibria, and the model would not be valid without an assumption on the observed selection of equilibrium. Because the network game is supermodular, there exist Pareto-best and -worst equilibria.[5] We assume that the world is characterized by some frequency of best and worst equilibria.

We estimate the parameters of the model using the method of simulated moments with indirect inference. We employ a simulation estimator because while our model is straightforward to solve for a given vector of parameters and draws on econometric unobservables, it would be extremely hard to analytically solve for the likelihood or conditional expectations of our dependent variables. One option would be to use simulated maximum likelihood. This can be problematic since the resulting estimator is not consistent for a fixed number of simulation draws. A more practical problem is that our endogenous variable is a mixed discrete-continuous variable – we observe banks either adopting or not adopting ACH, and conditional on adoption, we observe the number of transactions. Because of this, it would be hard to apply simulated likelihood techniques, since simulation will tend to generate probability 0 events. Using the method of simulated moments avoids these two problems, but because modeling correlations are so important to us (e.g. correlation in adoption decisions within market, correlation across time within bank, correlation across time within market), one quickly generates an unwieldy number of moments.

As a way of selecting important moments, we use indirect inference, which essentially specifies reduced-form regression parameters as moments. The first step is to run a set of regressions on the actual data. Then, given structural parameters, one simulates data and runs these same regressions on the simulated data. Estimates of the structural parameters are chosen to most closely match the regression estimates of the simulated data to those of the true data. For our

[5] See Milgrom and Shannon (1994).

model, the analysis of Gowrisankaran and Stavins (2004) provides a reasonable set of reduced-form regressions whose coefficients we use as indirect inference moments.

The remainder of this paper is divided as follows. Section 2 describes the model. Section 3 describes the data. In Section 4, we detail our estimation procedure, including the computation of the equilibrium and identification of the parameters. Section 5 contains results and Section 6 concludes.

2. Model

We propose a simple static two-sided market model of ACH adoption at a geographically local level. Consider a localized network of J banks in market m at time t, each with a given number of customers. The timing of our game is as follows. In the first stage, banks simultaneously decide whether or not to adopt the ACH technology. Let $A_{mt} = \left(A_{1mt}, \ldots, A_{Jmt}\right)$ be a set of indicator functions representing these adoption decisions. In the second stage, consumers decide whether to adopt ACH for their individual transaction originations. For a particular transaction (between two consumers) to be made through ACH, both consumers' banks must have adopted ACH, and in some cases which we detail below, both customers must have adopted the technology. Assume that all econometric unobservables are common knowledge to all firms and are unobservable only to the econometrician.

The timing of the game is as follows. Banks first simultaneously make their adoption decisions. Then, each consumer observes the adoption decision at her bank, and consumers simultaneously choose their adoption decisions.[6] We proceed by first analyzing consumer decisions conditional on A_{mt}. Then we move to the first stage and analyze equilibrium bank decisions.

[6] Because all econometric unobservables are common knowledge, consumers will be able to perfectly infer the equilibrium decisions of other banks and other consumers, even though these decisions are not directly observable.

Since ACH transactions are a small percentage of a bank's total business, we assume that the bank's consumer base and deposits are exogenous to our model of ACH usage. Denote the deposits under bank j's control as x_{jmt}. Assume that the number of total (both ACH and check) transactions that bank j's consumers engage in at time t is proportional to these deposits x_{jmt}, i.e.,

$$(1) \qquad T_{jmt} = \lambda x_{jmt}.$$

We assume that the demand for transactions is perfectly inelastic, and hence that prices of transactions do not enter into (1). We feel that this is a reasonable assumption because the demand for transactions is in fact likely to be fairly inelastic and because again ACH is a small proportion of total transactions.

While we assume that the total number of transactions that consumers make is a constant fraction of deposits, we do model the proportion of these transactions that are made through ACH, which we denote as T_{jmt}^{ACH}. We assume that each bank j has a set of consumers each of whom needs to originate N transactions in period t.[7] By definition, if bank j *has not* adopted ACH technology, these N transactions must be made through paper checks. If bank j *has* adopted ACH, the consumer does have the option of using ACH.

Consider consumer i's adoption decision conditional on her bank having adopted ACH. We assume that the consumer obtains net utility:

$$(2) \qquad V_{ACH} - V_{CHK} = \tilde{\beta}_1 + \tilde{\beta}_2 (p_t^{ACH} - p_t^{CHK})$$

[7] There are a number of dimensions in which this is a stylized model of consumer behavior – in particular the fact that consumers all make an identical number of transactions. This is necessary as we have no consumer level data on number of transaction originations.

from making an ACH (versus check) transaction, where p_t^{ACH} and p_t^{CHK} represent the prices of ACH and check transactions respectively. Note that prices do not vary cross-sectionally as they are set nationally by the Federal Reserve.

We assume that the consumer's transaction partners are allocated randomly among consumers of banks in the network, that the number of consumers is large enough to treat consumers as atoms, and that the net utility from an ACH transaction is positive. We allow for two types of ACH transactions. The first type are "two-way" transactions. "Two-way" ACH transactions require *both* originating and receiving customers to have adopted ACH; otherwise the transaction would be completed using a paper check. The second type are "one-way" transactions. "One-way" ACH transactions only require the *originating* customer of the transaction to have adopted ACH. These "one-way" transactions are intended to represent transactions such as payroll direct deposit or mortgage payments – where consumers do not formally need to adopt technology.[8] The type of transactions that we model as one-way are generally thought to form the bulk of ACH transactions. Note that while both consumers do not need to have adopted for a "one-way" transaction to occur, both the consumers' banks must have adopted. We measure the proportion of each consumer's total transactions which have the potential to be "two-way" ACH transactions with the parameter δ. $(1-\delta)$ is the proportion of potential "one-way" transactions.[9]

---

[8] In our model, adoption corresponds to the ability to originate transactions. Most recipients of direct deposit paychecks do not originate ACH transactions, and hence are not adopters in our sense. Note that both one- and two-way transactions can be debits or credits. A direct deposit for payroll is an originating credit transaction for the employer, while a mortgage payment is an originating debit transaction for the lender.

[9] An alternative model would specify two different types of consumers, one who only makes two-way transactions, and one who can make one-way transactions. Without individual level data, it would be hard to distinguish this model from our current model, where there is only one type of consumer, but where consumers make two types of transactions.

Given that the net utility from using ACH is assumed positive, any pair of consumers who have both adopted ACH will use ACH to process two-way transactions, and any consumer who has adopted ACH will use ACH for his one-way transactions with consumers at banks that have adopted ACH. Thus, the proportion of a consumer's transactions that will be made through ACH if he adopts is:

$$(3) \qquad u_{mt} = \delta \left[\frac{\sum_j A_{jmt} P_{jmt} x_{jmt}}{\sum_j x_{jmt}} \right] + (1-\delta) \left[\frac{\sum_j A_{jmt} x_{jmt}}{\sum_j x_{jmt}} \right],$$

where P_{jmt} is the proportion of consumers adopting at bank j in market m in time t. Note that the first term in brackets is the proportion of consumers that have adopted in market m at time t (at banks that have adopted). These are transaction partners with which two-way transactions can be made. The second term is the proportion of consumers that are at banks which have adopted ACH. These are transaction partners with which one-way transactions can be made.

Using the above definitions, we can write consumer i's net expected utility from adopting ACH (vs. not adopting) as

$$(4) \qquad EU_{ijmt} = N \cdot u_{mt} \cdot (V_{ACH} - V_{CHK}) + F_{ijmt}$$

where F_{ijmt} denotes the negative of the fixed costs of adopting. From (4), expected utility is the number of transactions that the consumer will make (N) times the proportion of those transactions that can be made through ACH (u_{mt}), times the utility gain from those ACH transactions ($V_{ACH} - V_{CHK}$) minus the fixed costs of adopting.[10]

[10] Note that in our model, consumers obtain marginal ACH utility solely from the N transactions that they originate, and not from transactions that they receive from other customers. We make the same assumption regarding banks. We could estimate an alternate specification where utility and profits are generated on both sides of the transaction.

In our empirical work, we want to allow for very general unobserved correlation in ACH transactions across markets, firms, and time, to separately identify the network benefits of ACH from differences in consumer fixed costs. To allow for this, we specify F_{ijmt} as:

$$F_{ijmt} = \beta_0 + \beta_3 t + \alpha_{jmt} + \varepsilon_{ijmt}, \tag{5}$$

where t is a time trend, β_0 and β_3 are parameters to estimate, α_{jmt} is a normally distributed bank level econometric unobservable, and ε_{ijmt} is an i.i.d. consumer level logit error. We then allow for a very general correlation structure of α_{jmt} - specifically, we let

$$\alpha_{jmt} = \alpha^A_{jmt} + \alpha^B_{jm} + \alpha^C_m + \alpha^D_{mt}, \tag{6}$$

where $\alpha^A_{jmt} \sim \text{iid } N(0,\sigma^2_{\alpha^A})$, $\alpha^B_{jm} \sim \text{iid } N\left(0,\sigma^2_{\alpha^B}\right)$, $\alpha^C_m \sim \text{iid } N\left(0,\sigma^2_{\alpha^C}\right)$, $\alpha^D_{mt} \sim \text{iid } N\left(0,\sigma^2_{\alpha^D}\right)$ and where α^A, α^B, α^C and α^D are all independent of each other.[11]

Substituting from (5) and (2) into (4), we obtain:

$$\begin{aligned} EU_{ijmt} &= \beta_0 + N \cdot u_{mt} \cdot \left(\tilde{\beta}_1 + \tilde{\beta}_2 (p^{ACH}_t - p^{CHK}_t)\right) + \beta_3 t + \alpha_{jmt} + \varepsilon_{ijmt} \\ &= \beta_0 + \beta_1 u_{mt} + \beta_2 p^{ACH}_t u_{mt} + \beta_3 t + \alpha_{jmt} + \varepsilon_{ijmt}, \end{aligned} \tag{7}$$

[11] As we detail the rest of the model, one might note that there are a number of places in the model where one might include a flexible unobservable structure like α_{jmt} in (6). This includes consumers' marginal benefits, consumers' fixed costs, banks' marginal profits, and banks' fixed costs). Because we essentially have one dependent variable in our analysis (number of ACH transactions), we felt that from an identification perspective it was only prudent to include one set of flexible unobservables. The reason we put them in consumer fixed costs is because this was the specification that appears to fit the data best.

where β_1 and β_2 are newly defined parameters, defined by $\beta_1 = N\cdot\left(\tilde{\beta}_1 - \tilde{\beta}_2 p_t^{CHK}\right)$ and $\beta_2 = N\cdot\tilde{\beta}_2$.[12] By integrating out over the logit error ε_{ijmt},[13] we get the proportion of consumers at bank j in market m in time t that adopt ACH as:

$$(8)\qquad P_{jmt} = A_{jmt}\frac{\exp\left(\beta_0 + (\beta_1 + \beta_2 p_t^{ACH})u_{mt}(P_{1mt},\ldots,P_{Jmt},A_{1mt},\ldots,A_{Jmt}) + \beta_3 t + \alpha_{jmt}\right)}{1+\exp\left(\beta_0 + (\beta_1 + \beta_2 p_t^{ACH})u_{mt}(P_{1mt},\ldots,P_{Jmt},A_{1mt},\ldots,A_{Jmt}) + \beta_3 t + \alpha_{jmt}\right)}.$$

Note that we have explicitly written out u_{mt} as depending on consumer adoption proportions and bank adoption decisions. In equilibrium, the vector of consumer adoption proportions at each bank, $(P_{1mt},\ldots,P_{Jmt})$ must satisfy the set of J equations defined by (8) conditional on bank adoption decisions $(A_{1mt},\ldots,A_{Jmt})$.

We next turn to optimal bank adoption decisions conditional on the above model of transaction choice. Recall that in the first stage, banks simultaneously decide whether to adopt ACH. Denote the marginal cost to the bank of an ACH and a check transaction as mc_t^{ACH} and mc_t^{CHK}, respectively. Assume that there is a per-period fixed cost FC of adopting ACH. Importantly, this is a per-period cost, not a one time sunk cost of adoption. As such, there are no dynamic optimization issues and firms simply maximize per-period profits.[14]

Banks compare profits from adopting ACH to profits from not adopting ACH. Given that bank j adopts, the total number of ACH transactions that its customers would originate is:

$$(9)\qquad T_{jmt}^{ACH} = T_{jmt}P_{jmt}u_{mt},$$

[12] We fold p_t^{CHK} into u_{mt} in (7) because we do not have data on the price of checks.

[13] Note that since we do not have consumer level data and include a flexible α_{jmt}, the assumption that the logit errors are iid is essentially WLOG.

[14] There is some evidence of this nature of fixed costs in our data as we see a number of banks switching from adoption to non-adoption between periods. See Gowrisankaran and Stavins (2004) for details.

This is the total number of transactions at bank j times the proportion of consumer adopters at bank j times the proportion of those consumers' transactions that can be made through ACH (both "one-way" and "two-way"). The increment in profits from adopting is the number of ACH transactions they would process times the difference in margins, minus the fixed cost of adoption. This increment is:

$$
\begin{aligned}
\Pi_{jmt}\left(T_{jmt}^{ACH}\right) &= T_{jmt}^{ACH}\left[\left(p_{jmt}^{ACH}-mc_{jmt}^{ACH}\right)-\left(p_{jmt}^{CHK}-mc_{jmt}^{CHK}\right)\right]-FC \\
&= T_{jmt}^{ACH}\times \text{markup}-\left(\overline{FC}+\alpha_{jmt}^{E}\right)
\end{aligned}
\tag{10}
$$

where fixed costs are divided into a common component ($\overline{FC}$) and an idiosyncratic component (α_{jmt}^{E}). As with the consumer fixed cost ε_{ijmt}, we normalize α_{jmt}^{E} to have a standard logistic distribution. We treat both $\overline{FC}$ and markup as parameters.

Bank j will adopt ACH at time t if and only if $\Pi_{jmt}\left(T_{jmt}^{ACH}\right)>0$. We can see that adoption will depend on other banks' decisions through T_{jmt}^{ACH}, which is a function of the equilibrium consumer adoption proportions $(P_{1mt},\ldots,P_{Jmt})$. An equilibrium $(A_{1mt},\ldots,A_{Jmt};P_{1mt},\ldots,P_{Jmt})$ requires that all banks' adoption decisions are optimal conditional on all other banks' adoption decisions, i.e.

$$
A_{jmt}=\left\{\Pi_{jmt}\left(T_{jmt}^{ACH}\left(A_{1mt},\ldots A_{j-1m,t},1,A_{j+1,mt},\ldots,A_{Jmt}\right)\right)>0\right\},\ \forall j,
\tag{11}
$$

where $T_{jmt}^{ACH}(\cdot)$ satisfies (3), (8) and (9).

Some customers in our model will have accounts at branches of banks whose headquarters are outside the network. We assume that these banks make their adoption decisions without considering the conditions in the network; i.e. their adoption decisions are exogenous to the unobservables in the market. However, conditional on adoption, customers at those banks choose their adoption decisions using the same criteria as banks whose headquarters are in the network. Thus, if a bank with headquarters outside the network chooses to adopt ACH, the probability of its customer adopting ACH will be given by (8). As in Gowrisankaran and Stavins (2004), these non-local banks will provide an important source of identification.

There can be multiple equilibria of this game if the magnitude of the network externalities are sufficiently large. To see this, note that on one hand, if every customer is using the network good, then any one customer is likely to want to use it. On the other hand, if no customer is using it, then any one customer is likely to not want to use it. This same logic is also true at the bank level. Because the value from another bank or customer adopting ACH is higher if the bank is itself adopting ACH, the adoption game is supermodular. Several properties follow from supermodularity.[15] These properties can easily be proved directly,[16] and do not depend on continuity but only on this monotonicity property. First, there exists at least one pure strategy subgame perfect equilibrium. Second, there exists one subgame perfect equilibrium that Pareto dominates all others and one (not necessarily distinct) subgame perfect equilibrium that is Pareto inferior to all others. Third, the proof of the second property is constructive, and it provides a very quick way to compute the Pareto-best and Pareto-worst subgame perfect equilibria. This last property is particularly important for estimation purposes.

To ensure an internally consistent specification, we need to specify the selection of equilibrium.[17] We estimate a specification that is consistent with the presence of multiple equilibria, and that can allow us to estimate whether markets tend to be in good or bad equilibria. Since we observe several separate networks, we allow for the possibility that some networks are in a good equilibrium while others are in a bad equilibrium. Specifically, we assume that there is some frequency that any given network is in the Pareto-best equilibrium, with a corresponding frequency of being in the Pareto-worst equilibrium. We estimate this frequency as a parameter. Formally, let $\omega_m \sim \text{iid } U(0,1)$. We assume that the market will be in the Pareto-best equilibrium

[15] See Milgrom and Shannon (1994).
[16] See Gowrisankaran and Stavins (2004).

if and only if $\omega_m < \Omega$, where Ω, the probability of being in the Pareto-best equilibrium, is a parameter that we estimate.[18] Note that we do not allow for the equilibrium to vary within a network across time.

3. Data

Our principal data set is the Federal Reserve's billing data that provides information on individual financial institutions that processed their ACH payments through Federal Reserve Banks.[19] We observe quarterly data on the number of transaction originations by bank for the period of 1995 Q2 through 1997 Q4. ACH transactions can be one of two types: credit or debit. A credit transaction is initiated by the payer; for instance, direct deposit of payroll is originated by the employer's bank, which transfers the money to the employee's bank account. A debit transaction is originated by the payee; for example, utility bill payments are originated by the utility's bank, which initiates the payment from the customer's bank account. For each financial institution in the data set, we have the ACH volume processed through the Federal Reserve each quarter.

We link these data with three other publicly available data sources. First, we use the Call Reports database, which provides quarterly information on deposits and zip code information for federally registered banks at the headquarters level. Second, we use the Summary of Deposits database, which provides annual information on zip code and deposits for banks at the branch level, in order to find small branches of non-local banks. Last, we use information from the census that provides the latitude and longitude of zip code centroids. Gowrisankaran and Stavins

[17] Heckman (1978) shows that this type of simultaneous equations model is not well-specified without some such assumption.

[18] Our method of estimating models with multiple equilibria is a generalization of the method used by Moro (2002) who estimates the equilibrium as a parameter. The difference is that we estimate the frequency of being in either equilibrium as a parameter, since we observe several regional markets, while Moro (2002) only has one market per year.

[19] We thank the Federal Reserve's Retail Payments Product Office for making this data set available to us.

(2004) provide more details of the data sources and linking. The resulting data set contains approximately 11,000 banks over 11 quarters.

Our estimation procedure is based on the assumption that a bank's network is geographically local. Our basic definition of a network is the set of banks whose headquarters are within 30 kilometers of the headquarters of a given bank. Because we are solving for an equilibrium of the adoption game, we need to also include all the banks that are within 30 kilometers of the banks that are within 30 kilometers, and all the banks that are near these banks, etc. We performed this process in order to separate our data set of 11,000 banks into mutually exclusive networks. Each network is self-contained, in the sense that every bank headquarters that is within 30 kilometers of any bank headquarters in the network is also in the network, and no bank headquarters in the network is within 30 kilometers of any bank headquarters outside the network.

One significant data problem is that many banks have become national in scope. As the relevant network for these banks is likely to be national, our model would not be particularly meaningful for these banks. Thus, we kept in our sample only banks that are in small markets. Specifically, we kept all networks with 10 or less bank headquarters total during every time period of our sample. From this set, we excluded networks where any one bank had more than 20 percent of its deposits outside the network, or where in aggregate, 10 percent of deposits for local banks were outside the network. We were left with a sample of 456 mutually exclusive networks comprising 878 local banks, observed over 11 time periods.

Figure 1 displays a map of New England with the networks from this region marked with asterisks and the major population centers marked with circles, in order to give some idea about typical networks. There are eight networks in New England in our sample, all of which are small, isolated towns, such as Colebrook, NH and Nantucket, MA. The major population centers, such as Boston, MA and Hartford, CT are all far away from these networks.

As described in Section 2, we use information from banks with branches in the network but headquarters outside the network, but model them separately from banks with headquarters in the network. We include in our sample 661 bank branches from banks outside the network.

Table 1 gives some specifics on the networks at every time period, broken down by the number of banks with headquarters in the networks. Approximately half of the network time-periods in our sample – 2730 in all – are composed of only one local bank. Another quarter of the network time-periods have two banks. However, there are large numbers of network time-periods with up to 10 local banks. Banks in our sample tend to be small banks, with assets of around $100 million. The percentage of firms using ACH appears to be quite consistent across network size, although banks in smaller networks have fewer ACH transactions.

Table 2 examines the non-local banks in these markets. Of note is the large number of outside banks. For instance, in markets with one local bank, the average number of outside banks is 2.74. Although the sizes of non-local bank branches and local banks are similar in terms of deposits, non-local banks adopt ACH much more frequently. This is due to the fact that the non-local banks are, on average, much larger than the local banks, and than their local branches.

Table 3 gives some specifics on the changes in ACH usage over our sample period. We can see that the fraction of banks using ACH increased during our sample period. Moreover, there appears to be a large fraction of networks where every bank uses ACH – more than one would expect without correlations in usage.

One factor that can affect usage of ACH is its price. Prices that the Federal Reserve charges banks for ACH processing are set at a fixed rate and adjusted periodically. Figure 2 displays a time series of these prices. Note that the intraregional per-item prices (that is, prices for ACH items exchanged between banks located within the same Federal Reserve District) did not change throughout our sample period. At the same time, the interregional prices declined from $0.014 in 1995 to $0.01 in 1997. In May 1997, the Federal Reserve implemented a two-tier price system of $0.009 for banks with less than 2500 transactions per file and $0.007 for banks with more than 2500 transactions per file. We ignore the $0.007 price because we do not have data on the number of transactions per file (only monthly totals) and because most of the banks in our sample are sufficiently small as to only pay the higher rate. Because prices are set by fiat and do not respond to changes in local demand, they may be viewed as exogenous. We do not have any information on the prices that banks charge to their customers. In addition to per-transaction costs, banks must file fees of \$1.75 per small file and \$6.75 per file per large file and pay an ACH participation fee of \$25 per month. Also, banks that offer ACH generally maintain a Fedline connection for ACH as well as other electronic payment services.

4. Estimation

Our model is based on a vector of unknown parameters $\left(\lambda,\beta,\Omega,\overline{FC},\text{markup},\sigma\right)$ and econometric unobservables $\left(\alpha,\omega\right)$.[20] For ease of notation, let us group the unknown parameters together as θ. Our estimation algorithm seeks to recover θ from the data. In this section, we describe this algorithm, including the computation of equilibria, and explain how the parameters of the model are identified.

4.1 Estimation Algorithm

We start by defining the data for one network in a given time period. For ease of notation, suppose that the network has J banks at every time period, that banks $1,\ldots,\hat{j}$ are local, and that banks $\hat{j}+1,\ldots,J$ are branches of non-local banks. For each bank, our data contain observed predetermined variables, namely its local deposits x_{jmt}, price p_t, time t, its local/non-local status and the adoption decisions of the branches of non-local banks. Let us group together the exogenous data as $z_m \equiv \left\{x_{1t},\ldots,x_{Jt},t,p_t,A_{\hat{j}+1,t},\ldots,A_{J,t}\right\}_{t=1}^{T}$. Our endogenous variables are the observed T_{jmt}^{ACH} for local banks only; note that $T_{jmt}^{ACH}>0$ implies that $A_{jt}=1$.

A natural way to proceed with estimation would be to construct the likelihood function for market m:

---

[20] Note that since there is a continuum of consumers, the consumer level unobservables ε are aggregated out of the model at the level of the data.

(12) $$L_m(\theta)=P\left(\left\{T_{t1}^{ACH},\ldots,T_{\hat{j}t}^{ACH}\right\}_{t=1}^{T}\left|z_m;\theta\right.\right).$$

Note that (12) treats an entire market over time as an observation. This is necessary because actions of banks (and consumers) are correlated within market and across time due to equilibrium behavior and correlated unobservables. The problem with this approach is that there is no closed-form solution for the density of the endogenous variables used in (12). This suggests the use of simulation methods. Many recent papers use simulated maximum likelihood methods to estimate structural models.[21] However, the observed decisions in these papers are typically discrete (e.g. replace a bus engine or not; go to school or work). In our case, our endogenous variable, the number of ACH transactions T_{jt}^{ACH}, is essentially continuous. While simulation methods can easily be used to evaluate a discrete probability, it is much more problematic, both numerically and theoretically, to use them to evaluate the density of a continuous random variable. The reason is that one will typically never generate simulation draws that lead to the observed outcomes, resulting in probability zero events.[22] Another problem with the simulated maximum likelihood approach is that it does not produce consistent estimates for a finite number of simulation draws, and our experimentation with this method revealed big differences in results with differences in the number of draws. (See also Keane (1994) and Ackerberg (1999) for evidence of the small sample bias this can create.)

We pursue an alternative approach, method of simulated moments (MSM). Consider the following moment:

(13) $$G(\theta)=E\left[y_m\text{-}E\left[y_m\left|z_m,\theta\right.\right]\left|z_m\right.\right],$$

[21] See Keane and Wolpin (2000) for instance.

[22] One alternative would be to use the methodology of Keane and Wolpin (2000). They add analytically integrable measurement error to the model, which generates a positive likelihood of any event.

where y_m is any function (or set of functions) of the observed endogenous variables for market m across all time periods, and $E\left[y_m \middle| z_m, \theta\right]$ is the expected value of that function. Since the observed data y_m was generated by the true parameter vector θ_0, $G(\theta)$ is by definition zero when evaluated at θ_0. The MSM approach replaces the non-analytically computable expectation $E\left[y_m \middle| z_m, \theta\right]$ with a simulated analog of the expectation. Define $\upsilon_s, s = 1, \ldots, S$ to be S simulation draws from the distribution of unobservables (α, ω) for every market. Given υ_s, we can easily (see section 4.2) compute the appropriate Nash equilibrium[23] of the industry in order to recover the endogenous variables. Denote by $y_m(z_m, \theta, \upsilon_s)$ the equilibrium endogenous variables given that draw. Replacing the inner expectation of (13) with its unbiased simulation analog $\frac{1}{S}\sum_s y_m(z_m, \theta, \upsilon_s)$, and the outer expectation with its data analog gives us:

$$(14) \qquad \hat{G}_M(\theta) = \frac{1}{M}\sum_{m=1}^{M}\left[\left(y_m - \frac{1}{S}\sum_s y_m(z_m, \theta, \upsilon_s)\right) \otimes f(z_m)\right],$$

Since the expectation of $\hat{G}_M(\theta)$ is zero for any choice of f, the MSM estimator that minimizes the norm of (14), i.e.

$$(15) \qquad \hat{\theta}_{MSM} = \arg\min_{\theta} \left\| \hat{G}_M(\theta) \right\|.$$

is a consistent estimator of θ even for a finite number of simulation draws (see McFadden (1989) and Pakes and Pollard (1989)). The logic behind this result is that the errors between the simulated expectation and the true expectation will average out over markets.

[23] Appropriate means Pareto-best or –worst depending on ω.

While the MSM estimator is consistent, there are two central problems that arise when we use it in the context of our model. First, there are many moments one could conceivably match. Our model contains a large number of endogenous variables. Since there are up to 10 banks per markets, there are up to 10 adoption and usage decisions per market per time period. All of these moment conditions should also be interacted with the predetermined variables to impose the conditional independence restrictions. Even more importantly, as we will infer network externalities from correlations in usage decisions, we would need to capture second- and higher-order cross moments; i.e. the usage of bank 7 times the adoption of bank 10 minus the expected value of this variable. To identify the random effects, we would then need to include similar moments across time (e.g. adoption of bank 1 in time 1 times adoption of bank 3 in time 5). We cannot possibly match all these moments. Not only is it computationally unwieldy (especially because there are a different number of banks in each market), but would also likely be non-negligible finite–sample biases with the large number of moments (see e.g. Staiger and Stock (1997)).

To address this first problem, we use a variant of the MSM estimator called indirect inference (II), proposed by Gouriéroux, Monfort and Renault (1993), which works as follows. First, suppose we perform some estimation routine on the true data, resulting in some parameter vector μ^{DATA}. This estimation routine may be "incorrect" in the sense that the estimated coefficients in the inference do not necessarily correspond to any structural parameters. As an example, we could use a reduced-form regression specified by Gowrisankaran and Stavins (2004), such as a regression of bank adoption on the fraction of other banks adopting, fixed effects and other controls. Second, for a given structural parameter vector θ' and simulation draw υ_s, we compute the equilibria of the model for each market, and perform the same "incorrect" estimation. Call the resulting parameter vector $\mu_s^{\theta'}$. Then, the II estimator is constructed as:

$$(16) \qquad \hat{\theta}_{II} \equiv \arg\min_{\theta} G^{II}(\theta) \equiv \arg\min_{\theta} \left\| \mu^{DATA} - \frac{1}{S}\sum_{s=1}^{S} \mu_s^{\theta} \right\|.$$

Thus, the II algorithm chooses the structural parameter vector for which the coefficients from the simulated data most closely match the coefficients from the actual data. With the II algorithm, one can sensibly match many fewer moments. For instance, the above reduced-form regression will capture the same insight as the standard MSM estimator – that with network externalities, usage levels for banks in a network should be correlated, even after conditioning on random effects – with many fewer moments. Also, note that even though the II estimator in (16) is not linear in each observation (as is (13)), Gouriéroux, Monfort and Renault (1993) show that the estimator will still be consistent for a fixed number of simulation draws. The result follows because the estimate for one simulation draw, μ_s^{θ} will converge in probability to $E\left[\mu^{\theta}\right]$ as the number of observations becomes large. In Section 4.3, we detail the exact choice of moments that we use in our II procedure.

The second problem with all MSM procedures, including the II variant is that there are infinitely many different norms $\|\cdot\|$ that can be used in (15) or (16) and different norms will give different estimators if there are more moment conditions than parameters to estimate. From the definition of a norm, any norm can be characterized by a weight matrix A, where A is a positive-definite matrix with dimension equal to the number of moment conditions. The "A" that will give efficient estimates is the inverse of the variance of the moment conditions in (16), evaluated at the true parameter value. For standard MSM, the typical way to obtain asymptotically efficient estimates is to find first-stage consistent estimates by using any weight matrix, and then to approximate the efficient weight matrix by finding the variance of the moment conditions evaluated at the first-stage estimates. Unfortunately, this can be problematic in practice as first–stage estimates can vary greatly with the initial weight matrix. The II estimator has an advantage in that one can compute an asymptotically efficient weight matrix in the first stage without relying on structural parameter estimates. The initial estimation of the coefficients μ^{DATA} will typically generate a variance matrix of the parameters, Var(μ^{DATA}). Since, according to the model, each μ_s^{θ} has the same distribution as μ^{DATA} (at the true θ), the variance of $\mu^{\mathrm{DATA}} - \frac{1}{S}\sum_{s=1}^{S}\mu_s^{\theta}$ is simply $\left(1+\frac{1}{S}\right)\mathrm{Var}(\mu^{\mathrm{DATA}})$. We use the inverse of this for our weight matrix, and there is no need to perform a two–stage estimation process. The underlying reason that we can find this weight matrix is that none of the II moments are conditional on exogenous regressors, making it very easy to find the variance of the moments.

The only caveat in our particular case is that even though μ^{DATA} is composed of regression coefficients, we cannot use the standard OLS variance/covariance matrix because the regression unit of observation is a bank and we expect clustering of errors at the market level. In addition, we run multiple reduced–form estimations for our indirect inference and OLS variance/covariance matrices would not give us covariances between coefficients across the regressions. To address these issues, we use bootstrap methods to compute Var(μ^{DATA}). We resample the data with replacement, and calculate μ^{DATA} for 3,000 resampled data sets in order to numerically approximate its variance. To consistently resample given that we allow for random effects that link observations across time, our unit of observation for the sampling process is one market over time.

Var(μ^{DATA}) is also important for computing the standard errors of our estimated parameters. From Gouriéroux, Monfort and Renault (1993), the asymptotic variance matrix is given by:

$$\mathrm{Var}(\theta) = \left(\Gamma' A \Gamma\right)^{-1}$$

where $\Gamma = \frac{\partial G^{II}}{\partial \theta}$ and A= $\left[\left(1+\frac{1}{S}\right)\mathrm{Var}(\mu^{\mathrm{DATA}})\right]^{-1}$ is the weight matrix defined above.

4.2 Computation of Equilibrium

In order to compute our estimator, we need to compute simulated equilibria for each simulation draw. This involves solving for the Pareto-best or -worst subgame perfect equilibrium of the model conditional on a vector of pre-determined variables and econometric unobservables.

In general, estimation of Nash equilibria can be very computationally intensive. This computational intensity is a large part of the reason why structural models are notoriously difficult to estimate. In our case, it is computationally simple to solve for both subgame perfect equilibria. The underlying reason for this is that the network externality is assumed to always be positive, which makes the game supermodular. Because of this, the optimal reaction functions will always be a monotone mapping of the previous stage reaction functions. This is also the basis of the proof that there is a Pareto-best subgame perfect equilibrium given in Gowrisankaran and Stavins (2004), Proposition 1.

Thus, we solve for the Pareto-best subgame perfect equilibrium by using the following iterative process on adoption of banks and consumers. We start the first iteration by assuming that all banks and consumers use ACH, i.e. $\left(A^1_{1mt}=1,\ldots,A^1_{Jmt}=1\right)$ and $\left(P^1_{1mt}=1,\ldots,P^1_{Jmt}=1\right)$. In the second iteration we consider each bank in turn. For bank j, we find the consumer adoption decisions given the adoption decisions in the first iteration, except with the assumption that bank j has adopted ACH.[24] We then determine whether bank j would find it profitable to adopt given this level of usage, and enter this as the new strategy. We repeat this process for each bank. This results in a vector $\left(A^2_{1mt},\ldots,A^2_{Jmt},P^2_{1mt},\ldots,P^2_{Jmt}\right)$ where each level is weakly less than in the first iteration. We repeat this process until convergence; convergence is guaranteed by this monotonicity property. As in Gowrisankaran and Stavins (2004), we can show that the limiting values $\left(A^N_{1mt},\ldots,A^N_{Jmt},P^N_{1mt},\ldots,P^N_{Jmt}\right)$ form a Pareto-best subgame perfect equilibrium. Correspondingly, if we start the first iteration by assuming that no one is using ACH, i.e. $\left(A^1_{1mt}=0,\ldots,A^1_{Jmt}=0\right)$ and $\left(P^1_{1mt}=0,\ldots,P^1_{Jmt}=0\right)$ and then iterate to convergence, the algorithm will converge to the Pareto-worst Nash equilibrium.

[24] Recall that consumers of bank j observe the decisions of bank j before making their adoption decisions.

We can also use variants of this algorithm to solve for the outcomes when local banks internalize the network externality and when consumers internalize the externality, both of which we report. For the case of banks internalizing the externality, we solve for the bank adoption decisions differently, assuming that banks value the difference in profits from *all* banks resulting from their adoption decision. For the case of the consumers internalizing the externality, we solve for the optimal cutoff fixed cost for each consumer, which differs from the non-cooperative case, even conditional on other agents' actions.

Because of the monotonicity of the reaction functions, our algorithm converges to the appropriate Nash equilibrium very quickly. For instance, to evaluate one parameter iteration with 10 simulation draws, we require computing a Nash equilibrium for the roughly 500 markets over 11 time periods with 10 different simulation draws and 2 equilibria. It takes about 3 seconds to solve for these 100,000 equilibria on a modern workstation. The reduced-form II moments, which include fixed effects for every bank, take much longer to compute, approximately 30 seconds per vector of simulation draws, or 5 minutes with 10 simulation draws.

4.3 Identification and Choice of Indirect Inference Moments

In this subsection, we discuss the exact choice of moments that we match and explain how they affect the identification of the parameters. Recall that because we use II, our moment conditions are the differences in the coefficients of the indirect inference estimation procedure on the actual data and the same indirect inference estimation on the simulated data. Thus, it suffices to describe these estimations.

While the choice of II moments might be difficult in general, Gowrisankaran and Stavins (2004) provide a set of three reduced-form analyses that are useful in determining the presence of network externalities for this market. To capture the same sources of identification, we use essentially the same regressions, although we choose linear functional forms to reduce the computational time since the exact functional form is not important. We also cannot use every coefficient from these regressions as moments, since they include thousands of coefficients (because of fixed effects). Hence, we use as moments the coefficients that the earlier paper found to be indicative of network externalities, as well as other coefficients to normalize the scale of the predicted values to the scale from the data.

The first method examines clustering of adoption decisions within a network. We regress the adoption decision of banks on the deposit-weighted adoption decisions of competitors' banks, deposits and squared deposits and bank and time fixed effects. We use the 3 main regression coefficients and 11 time dummies as moments, as well as the standard error of the coefficient on competitor adoption. We also decompose the residual and the fixed effects into market– and firm–specific effects, and use the four resulting standard deviations as additional moments. As in Gowrisankaran and Stavins (2004), we also regress ACH volume per transaction on competitors' volume per transaction and deposit-weighted adoption decisions, using the same deposit and fixed effect controls. We take three moments from this regression: one of the time dummies (capturing the constant term), and the volume and adoption measures.

The second method is to regress adoption on a Hirschmann-Herfindahl Index (HHI), time dummies and 100 indicators for deposit size, ranging from 0 – \$2 million, \$2 – 4 million, etc. We use two coefficients, one time dummy and the HHI, and the R^2 (a summary statistic for the variance of the residuals) as moments here.[25]

The third method is to regress local bank adoption decisions on the adoption decisions of non-local banks. We create a variable that indicates the deposit-weighted fraction of non-local banks that adopt. We regress local bank adoption on this measure, deposits and squared deposits squared and time dummies. We chose as moments the coefficient on the non-local adoption measure, one of the time dummies and the R^2 from the regression. We then redefined the measure to be one in the case of no non-local banks, and reestimated the model, taking the same three moments for this new regression.[26] We have a grand total of 31 moments, all listed in Table 5.

[25] We include deposits categorically, rather than linearly, for this regression, to avoid the opposite and confounding effect of economies of scale.

[26] The logic behind this is that banks without any non-local competitors are more similar to banks where all non-local competitors have adopted, then to ones where none have adopted.

In our model, the network effects are captured by five parameters: the consumer and bank fixed costs of adoption, the consumer and bank per-transaction benefits from adoption and the probability of two-way transactions. To understand the identification of the parameters, note that if we knew the marginal benefits of ACH adoption (for banks the relative markup for ACH, for consumers the relative utility from an ACH transaction) and the probability of two-way transactions, the levels of adoption decisions would identify the two fixed costs of adoption. In other words, the observed proportion of bank adoption would identify $\overline{\mathrm{FC}}$, and the observed proportion of consumer decisions would identify β_0.[27] Thus, the key concern is how the parameters on the marginal benefits and two-way transactions are identified.

Our model identifies these parameters via three separate mechanisms. The first source of identification is covariance restrictions. We assume that after controlling for bank and market characteristics with random effects, unobservables affecting adoption are independently distributed across banks in a given market. Thus, the estimation will find network externalities from this source if, after controlling for the random effects, the pattern of adoption within a network displays correlations consistent with network externalities. The second source is exclusion restrictions, based on the fact that the sizes of other banks do not enter into a bank's adoption decision. The estimation will find network externalities from this source if, for example, concentrated markets experience more ACH adoption. The third source of identification is the variation in adoption decisions by large, non-local banks. We assume that the adoption decisions of these banks are exogenous, and not made in response to equilibrium conditions in the market, but allow the customers at that bank to make their usage decisions in equilibrium. Our formal model of equilibrium allows us to combine all these sources of identification into one estimation procedure, and our II approach allows this to be done in a very transparent way.

One difference between the identification of this model and of Gowrisankaran and Stavins (2004) is that we have to jointly identify three crucial parameters (benefit parameters for

[27] Note that there is a selection issue here, since we do not observe the proportion of consumers adopting for banks

banks and consumers and the two-way transaction parameter) instead of just one. We can identify these parameters looking at both consumer adoption and bank adoption. As an example, consider the identification from the adoption of large, non-local banks. As the adoption of non-local banks exogenously increases, the equilibrium adoption rate of local banks should increase. The extent of this increase should identify the marginal benefit parameter for banks. Even conditional on local bank adoption (or conditional on local banks that were already adopting), the exogenous increase in adoption of non-local banks will also increase the probability that consumers of the local banks adopt, and thus increase the number of transactions. The shape of the increase in number of transactions will differ based on the values of the consumer benefits and two-way transactions parameters. For instance, with all two-way transactions and very large consumer benefits, the number of transactions will increase as the square of bank adoption (e.g. if one third of banks adopt and all of their consumers adopt, then one-ninth of transactions will occur between two consumers that have adopted. With large consumer fixed costs, the adoption curve will be more convex. With more one-way transactions, it will be closer to linear.

Another interesting identification issue is the equilibrium selection parameter Ω. Ω should be identified by differences in usage given different industry structures. For instance, as the number of firms increases, the increasing externality should make it more likely that there is a Pareto-worst equilibrium that is distinct from the Pareto-best equilibrium. Thus, we can identify the equilibrium selection parameters by examining whether there is increased unexplained variance in behavior for networks with more than one bank that does not exist for networks with one bank. Note that if we saw a high variance in the usage levels in *all* markets, this could be evidence of high variances of the random effects α, not necessarily multiple equilibria. Note that the polar cases of always in a good equilibrium and always in a bad equilibrium will be hard to separately identify.

5. Results and Implications

that do not adopt.

Using the indirect inference procedure developed in Section 4, we have estimated structural parameters of our model. We first present the results and then present policy experiments.

5.1 Results

Table 4 reports our estimates, while Table 5 exhibits the 31 indirect inference moments that were fitted. Examining Table 5 suggests that our model fits the data reasonably well. Comparing the differences in the simulated coefficients and the data coefficients to the bootstrapped standard errors of the coefficients provides some indication of this. Of the 31 marginal differences, only 5 are significantly different from zero. The moment condition at the estimated parameters is 53.51. As our choice of weighting matrix standardizes the moments, this represents the sum of 31 squared i.i.d. N(0,1) random variables. While a Chi-squared test rejects (barely) the joint hypothesis that all the differences are zero (at 99% confidence), it is not unusual for a structural model to be rejected by the data.

Turning to the actual estimates in Table 4, the parameters seem reasonable. For instance, the coefficient on time trend is positive, suggesting that there is increased acceptance of technological goods over time. The ACH price coefficient is negative. On the consumer side, both consumer fixed costs and marginal benefits are positive (consumer fixed cost = $-\beta_0$) and the ratio appears reasonable. On the other hand, for banks, the estimated mean fixed costs of adoption seems small in comparison to the normalized net markup (although the markup is per 1000 ACH transactions). This is consistent with the small bank–level fixed costs noted in Section 3. The relative magnitudes of these parameters are easier to interpret when we examine the economic significance of the parameters below.

At the bottom of the table is our estimated value of δ, the proportion of two-way transactions. We obtain δ=0, suggesting that *all* transactions are one-way transactions. This is an interesting result, suggesting that ACH is mainly being used by big companies (employers, mortgage companies) who have the ability to make ACH transactions with smaller consumers who have not formally adopted originating technology. Since one-way ACH transactions can be performed without both consumers adopting, this estimate suggests that there are no significant externalities *between* consumers.[28] It is interesting to see what in the data is generating our estimate of δ=0. As described in section 4.3, identification of δ should come from consumer adoption rates conditional on bank adoption rates. A simple regression of (ACH transactions per dollar of deposits) on (fraction of banks adopting (deposit weighted)) shows a linear relationship between the two variables.[29] This suggests that as more banks adopt, transactions go up proportionally. This is not consistent with two-way transactions, where transactions would go up more than proportionally. While the standard error of this parameter is relatively small (e.g. we can easily reject that all transactions are two-way, i.e. δ=1), the estimated parameter is on the boundary of parameter space, making its estimated standard error not completely reliable.

The estimates of the four random effects standard deviations show that the estimated standard deviations of α^{A}_{jmt} and α^{B}_{jm} are considerably higher than those of α^{C}_{m} and α^{D}_{mt}, suggesting that firm–specific effects are quantitatively more important than the market–specific effects. Our estimated value of the equilibrium selection parameter Ω is 0, suggesting that we are always in the Pareto-worst equilibrium. However, as will be clear in Table 6, there is no evidence of economically significant multiple equilibria at our estimated parameters. Given the lack of two distinct equilibria, Ω is not well–identified, and hence the estimate is not really indicative that we are stuck in a bad equilibrium. This is also indicated by the high standard error on the estimate of Ω, although again the estimated parameter is on the boundary of parameter space.

Although the parameter estimates are interesting in of themselves, it is much more valuable to examine the impact of the parameters on the estimated equilibrium. This is done in

[28] If we estimated the model where recipients of transactions obtain utility (or profits) (see footnote 11), there would be externalities between consumers, even though the proportion of two-way transactions would be zero.

Table 6. We look at 3 statistics of the estimated equilibrium – the percentage of local banks adopting ACH, the percentage of consumers adopting, and percentage of overall transactions done through ACH. At the estimated parameters, 68.4% of local banks are adopting, 17.8% of consumers are adopting, and 16.2% of all transactions are ACH.

The second row of Table 6 examines what our model predicts if there were no mean bank fixed costs of adoption. The difference between this and the first row is indicative of the level of the network externalities at the bank level. Although many more banks adopt ACH, the differences in consumer adoption rates and in transactions processed with ACH are small. This is due to our small estimated bank fixed cost of adoption, which implies that the holdup from consumers not using ACH is not due to their banks. On the other hand, when we eliminate the consumer mean fixed cost of adoption, there are big changes in the equilibrium proportion of consumers that adopt ACH: consumer adoption increases to 55.5%.[30] In response to this expected adoption by consumers, banks also increase adoption, to 96.8%, and in this equilibrium, 52.8% of all transactions are done using ACH. These estimates suggest that consumer fixed costs are the primary impediments to ACH adoption.

The next two rows of Table 6 examine the existence of multiple equilibria at our estimated parameter values by forcing either the Pareto-worst or the Pareto-best equilibria. The results across the two equilibria are very similar, though not identical. This suggests that at our estimated parameters, multiple equilibria are not a significant issue.

Last, we investigate what would happen if some of these externalities could be internalized. There is no natural way to compare consumer utility to firm profits. However, we can investigate what happens if all the local banks coordinated decisions to maximize joint profits, or if all consumers coordinate to maximize joint utility. Results are in the last two rows of Table 6. Joint profit maximization of all the local banks raises adoption, but not by much.

[29] In other words, higher order terms are insignificant. For example when we use fraction of banks adopting and fraction of banks adopting squared as explanatory variables, we get a T-statistic of 3.04 on the linear term and 0.18 on the squared term (the constant term is also insignificant).

[30] This is less than 100% because we are only eliminating the mean fixed cost of adoption; consumers with fixed costs higher than the mean may still not adopt. The same is true in the previous experiment where we eliminated bank *mean* fixed costs.

When all consumers coordinate to maximize joint utility, nothing changes. This results follows directly from the fact that all transactions are one-way, i.e. that $\delta=0$. Because of this, consumers simply do not exert externalities on each other. However, as we will see in the next section, consumers are exerting externalities on banks.

5.2 Policy Experiments

These results suggest that it is consumer fixed costs that are preventing widespread adoption of ACH technology. In contrast, bank fixed costs are small and do not significantly limit ACH use. This suggests that government policy, particularly at the consumer level, might increase welfare. We examine this possibility in Table 7.

The first column of Table 7 again examines properties of the estimated equilibrium. In addition to statistics on consumer and bank adoption, we report welfare measures – the sum of firm profits and the sum of consumer utilities. We have no way of converting these measures into dollars, so it is important to realize that these measures are not comparable to each other. Consumer utility is measured in "utils", and profits are measured in "profit units."

The second two columns essentially repeat two of the experiments of the prior section. We remove, sequentially, consumer and bank mean fixed costs through a government subsidy. Rows 6 and 7 of the table report the cost to the government (in profit units and utils respectively) of these policies. Rows 8 and 9 report the total profit units (bank profits – government cost in profit units) and total utils (consumer utils – government cost in utils) resulting from these policies respectively. With the bank subsidy, it is again clear that bank fixed costs are simply not large enough to prevent adoption. The consumer subsidy is far more effective at increasing ACH usage. Also note the extremely large benefits to banks from this consumer subsidy, as they are able to make considerably more variable profits. However, the subsidy is inefficient in terms of utils, as the gain in consumer utility due to the policy is more than offset by the cost of the subsidy to the government in utils. This suggests that the subsidy might be generating inefficiently high levels of adoption. Interestingly, although the bank subsidy does not increase total profits by much, it does unambiguously increase welfare, as both total profits and total utils go up.

Considering a subsidy of mean fixed costs is rather arbitrary, as there are distributions of these fixed costs at both the bank and consumer level. Columns 4 and 5 consider very large subsidies to banks and consumers, subsidies large enough to get virtually everyone to adopt (note that since there are some non-local banks who do not adopt, we cannot get all consumers to adopt). Because these subsidies are so large, one should not pay much attention to the Firm Profit and Cost to Gov't (in profit units) numbers in the very large bank subsidy case or the Consumer Utility and Cost to Gov't (in utils) numbers in the very large consumer subsidy case. What is more relevant are the Total Utility and Total Profit numbers in the last two rows of the table. While the very large consumer subsidy clearly generates lots of ACH transactions and lots of firm profits, there are extremely limited changes with the bank subsidies.

It is hard to make any conclusive evaluations regarding the above policies. This is because in almost all cases, either total utils or total profits go down as a result of the policy. Since we have no way of relating the increases in profits to the decrease in utils (or vice-versa), we cannot conclude that these policies are welfare improving. Recall that with the mean consumer fixed cost subsidy, inefficiently high levels of adoption appear to occur. With a smaller, more efficient, consumer subsidy, we might hope to keep even or increase total utility (as well as increasing total profits). Column 7 exhibits results from the largest consumer subsidy (approximately) that does this, 17% of their mean fixed cost. With this subsidy, total utils are unchanged, but firm (and total) profits increase by more than 33%. Column 8 adds a possible firm subsidy to the policy, noting from above that firm subsidies, while not increasing adoption by much, do unambiguously increase welfare. With a concurrent large bank subsidy, we can increase the consumer subsidy to 31% of their mean fixed costs and still keep consumer utility at its baseline level. This ends up increasing firm profits by about 75%. This is a pretty significant increase in welfare as a result of a simple subsidization policy.

One last interesting point is how our estimate that δ=0 impacts externalities. As noted in Section 5.1, when all transactions are one-way, consumers do not generate externalities towards other consumers. Given our estimates of bank fixed costs are fairly low, it appears that banks are not exerting large externalities on either other banks or consumers. So what is the main externality occurring in our data? It is the externalities that consumers impose on banks. This is evidenced by the 2nd and 4th columns of Table 7. With medium or large subsidies to consumers that substantially increase consumer adoption, bank profits increase 400-800%. This is a particularly interesting result since, at least in theory, one might think this externality could be internalized through subsidies in prices by banks to consumers.[31] Note that our estimated model is completely consistent with banks subsidizing (their own) consumers – our estimated fixed costs and marginal benefits (for both banks and consumers) just need to be interpreted as post-subsidy costs and benefits. In any case, it seems interesting that banks are not able to mitigate this externality through subsidies. Perhaps they *are* in fact subsidizing consumers, but not enough to completely eliminate the externalities (in other words, if there were no subsidization going on, we would have estimated much larger externalities). This coincides with some casual evidence on the issue. First, it appears that many ACH transactions are subsidized to negative prices – e.g. free checking for direct deposit. Perhaps banks are uncomfortable offering even more negative prices. Second, from conversations with small businesses, there appear to still be substantial markups for business-to-business transactions.

6. Conclusions

In this paper, we have estimated a structural equilibrium model of network externalities in the ACH banking industry in order to estimate the causes and magnitudes of network externalities for this industry. Our parameter estimates are reasonable, generally precisely estimated, and fit the data reasonably well.

[31] Since our assumption is that only originating banks profit from ACH transactions, these consumer externalities only affect their *own* banks. This makes it even more likely that the externality could be internalized.

We find that bank fixed costs from ACH adoption are low and do not explain why ACH is not more widely used. In contrast, consumer fixed costs of ACH adoption are substantial, and are a major explanation for the lack of ACH usage. Most ACH transactions appear to be one-way in the sense that bank adoption, but not consumer adoption, is necessary for them to be realized. Changes that lower the consumer fixed cost of ACH adoption will encourage adoption and usage of ACH. As electronic payment technologies become more widely accepted and used at the consumer level, we will expect to use vastly more ACH transactions.

Although we estimate that the Pareto-worst equilibrium is not identical to the Pareto-best equilibrium, we find that the two equilibria are very similar to each other in their implied ACH adoption decisions. Because the bank fixed costs are so low, the equilibrium bank ACH adoption is very close to the first best adoption level. Policies that subsidize a small portion of consumer fixed costs can unambiguously increase total surplus. Large bank subsidies for ACH adoption in conjunction with small consumer subsidies increase welfare even more.

References

Ackerberg, D. (1999) "An MSM Estimator for Discrete Choice Panel Data with Arbitrary Error Structure" mimeo, Boston University

Angrist, Joshua D. and Guido Imbens W. (1994). "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62: 467-76.

Brock, William A. and Steven N. Durlauf (2001). "Discrete Choice with Social Interactions." *Review of Economic Studies* 68: 235-260.

Gouriéroux, Monfort and Renault (1993) "Indirect Inference" *Journal of Applied Econometrics* 8:85-118

Gowrisankaran, Gautam and Joanna Stavins (2004), "Network Externalities and Technology Adoption: Lessons From Electronic Payments," forthcoming, *RAND Journal of Economics*.

Heckman, James J. (1978), "Dummy Endogenous Variables in a Simultaneous Equation System," *Econometrica* 46, 931-959.

Heckman and Robb (1985). "Alternative Methods for Evaluating the Impact of Interventions." In Longitudinal Analysis of Labor Market Data ed. by James J. Heckman and Burton Singer. New York: Cambridge University Press.

Keane, Michael (1994) "A Computationally Practical Simulator for Panel Data" *Econometrica*, 62 95-116

Keane, Michael P. and Kenneth I. Wolpin (1997), "The Career Decisions of Young Men," *Journal of Political Economy* 105, 473-522.

Keane, Michael and Kenneth Wolpin (2000). 'Estimating the Effects of Welfare on the Education, Employment, Fertility and Marriage Decisions of Women." Mimeo, Yale University.

Manski, Charles (1993). "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies* 60: 531-42.

Milgrom, Paul and Chris Shannon (1994). "Monotone Comparative Statics," *Econometrica* 62: 157-180.

Moro, Andrea (2003). "The Effect of Statistical Discrimination on Black - White Wage Inequality: Estimating a Model With Multiple Equilibria." *International Economic Review* 44: 467-500.

Rust, John (1987). "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." Econometrica 55: 999-1033.

Rysman, Marc (2004). "Competition Between Networks: A Study of the Market for Yellow Pages." forthcoming, *Review of Economic Studies*.

Topa, Giorgio (2001). "Social Interactions, Local Spillovers and Unemployment." *Review of Economic Studies* 68: 261-296.

Table 1: Characteristics of Banks in Network

Number of banks based in network	Number of networks/time periods	Mean deposits	Mean percent of banks using ACH	Mean ACH transactions by bank
1	2730	$45.8 Mil.	64.3%	457.7
2	1310	$49.5 Mil.	64.5%	452.0
3	367	$59.4 Mil.	67.8%	1217
4	172	$73.0 Mil.	74.4%	1348
5	83	$50.1 Mil.	74.2%	912.5
6	51	$125 Mil.	70.3%	3485
7	31	$139 Mil.	73.7%	2155
8	41	$57.5 Mil.	66.2%	991.5
9	39	$79.9 Mil.	69.5%	897.9
10	25	$81.6 Mil.	57.6%	732.2

Table 2: Characteristics of Branches of Non-Local Banks

Number of banks based in network	Mean number of non-local banks	Std. dev. of number of non-local banks	Mean deposits within network by non-local banks	Mean total deposits by non-local banks	Percent of non-local banks using ACH
1	3.43	2.74	$59.8 Mil.	$10.4 Bil.	88.5%
2	2.50	2.38	$60.2 Mil.	$6.8 Bil.	85.8%
3	4.05	3.21	$96.0 Mil.	$9.2 Bil.	89.0%
4	4.34	3.32	$92.0 Mil.	$8.6 Bil.	88.5%
5	6.15	5.16	$187 Mil.	$4.8 Bil.	83.3%
6	5.67	5.20	$96.9 Mil.	$8.5 Bil.	84.1%
7	9.13	4.26	$78.6 Mil.	$5.0 Bil.	91.9%
8	6.80	4.65	$95.2 Mil.	$7.9 Bil.	86.0%
9	8.72	5.80	$104 Mil.	$6.9 Bil.	87.4%
10	6.56	3.80	$120 Mil.	$4.9 Bil.	81.1%

Note: Table based on observations kept in sample.

Table 3: Usage Over Time by Banks in Network

Time Period	# of networks with no firm using ACH	# of networks with some, but not all, firms using ACH	# of networks with all firms using ACH
1995: Q2	14.3%	57.1%	28.6%
1995: Q3	16.8%	57.4%	25.7%
1995: Q4	17.3%	55.8%	26.9%
1996: Q1	14.3%	55.6%	30.1%
1996: Q2	10.9%	51.6%	37.5%
1996: Q3	12.5%	51.0%	36.5%
1996: Q4	8.4%	50.3%	41.4%
1997: Q1	7.3%	46.1%	46.6%
1997: Q2	5.8%	41.3%	52.9%
1997: Q3	7.1%	42.9%	50.0%
1997: Q4	6.1%	42.2%	51.7%

Note: Table includes networks with 2 or more banks kept in sample.

Table 4: Parameter Estimates

Parameter	Value	Standard Error
λ (transactions coefficient)	3.436	1.478
β_0 (consumer fixed benefit)	-3.050	0.172
β_1 (consumer marginal benefit)	0.825	0.213
β_2 (price coefficient)	-0.542	0.101
β_3 (time coefficient)	0.085	0.005
Markup	190.445	77.998
$\overline{FC}$ (bank fixed costs)	6.266	0.754
Ω (Probability of good equilibrium)	0	5.762
σ_{α^A} (std. dev. of random effect α^A_{jmt})	0.830	0.060
σ_{α^B} (std. dev. of random effect α^B_{jm})	1.953	0.132
σ_{α^C} (std. dev. of random effect α^C_{m})	0.129	0.062
σ_{α^D} (std. dev. of random effect α^D_{mt})	0.036	0.022
δ (proportion of two way transactions)	0	0.111
Moment condition at estimated parameters (31 moments)	53.31	

Table 5: Indirect Inference Moments at Estimated Parameters

Regression	Description	Moment in data	Moment in model	T-statistic for difference
Adoption on competitor adoption	Time dummy	.651649	.561715	-2.64828
	Competitor adoption	.0894093	.197719	3.07
	Time dummy	.0040079	.0090186	.883181
	Time dummy	.0173146	.0291377	1.42804
	Time dummy	.035803	.0450925	1.05796
	Time dummy	.0468305	.0515382	.495468
	Time dummy	.0642336	.0681886	.356067
	Time dummy	.0832865	.0759623	-.657472
	Time dummy	.089917	.0837702	-.517556
	Time dummy	.111271	.0959215	-1.21191
	Time dummy	.11248	.10549	-.577643
	Time dummy	.114206	.110365	-.31133
	Deposits	.147971	.312019	.857063
	Squared deposits	-.0879333	-.153278	-.257447
	SE competitor adoption	.0122277	.0119757	-.798558
	Std. dev. of residual	.197807	.170946	-2.45925
	Std. dev. of residual	.282577	.28912	.942498
	Std. dev. of residual	.123224	.112591	-1.96097

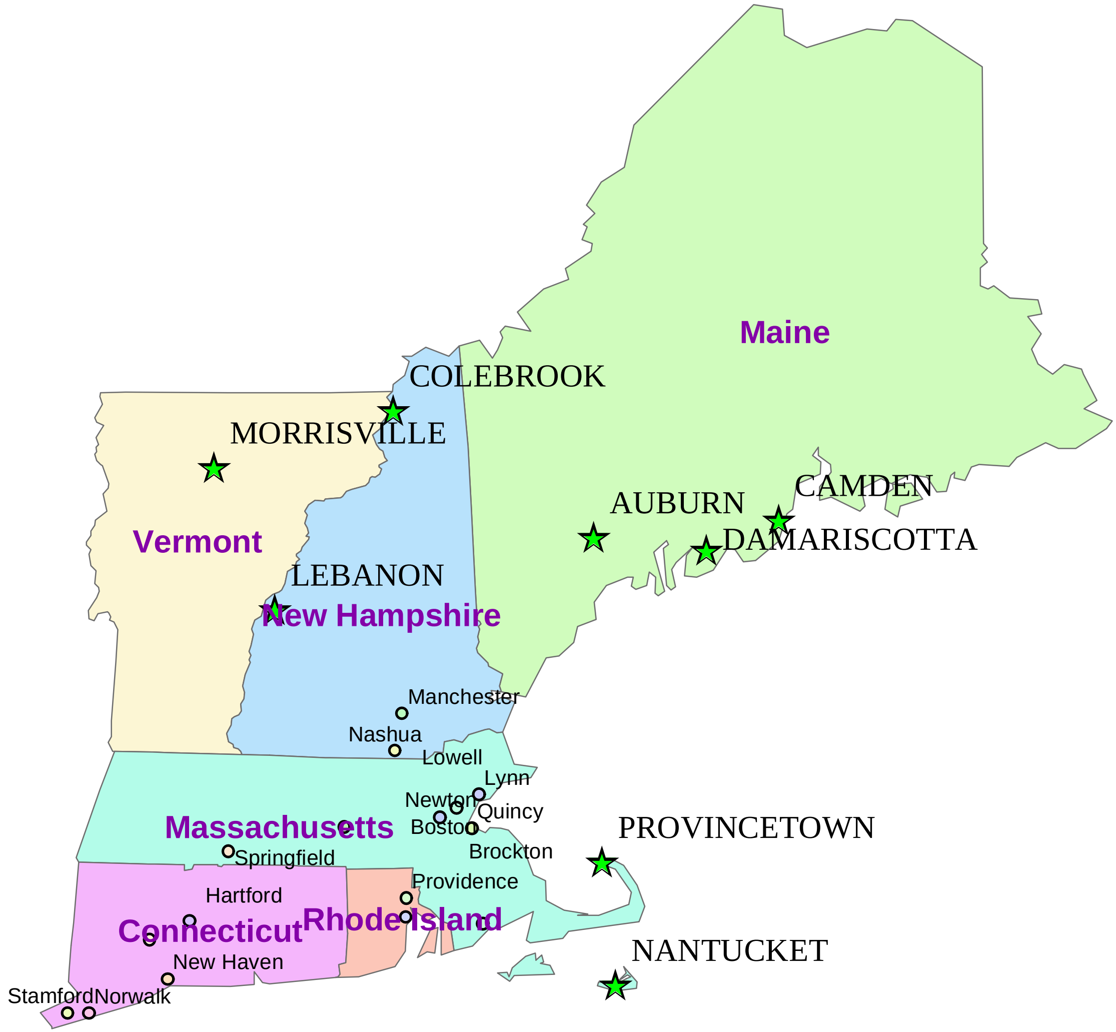
	Std. dev. of residual	.216824	.217475	.12613
ACH volume per deposits on competitors' volume and adoption	Time dummy	.0009091	.0001846	-.754543
	Competitor volume per transaction	.0020746	.0578511	1.44821
	Competitor adoption	.0000165	.0002584	2.17627
Adoption on HHI	Constant	.178407	-.0979803	-1.10996
	HHI	.0949221	.0691279	-.475707
	R^2	.136031	.112621	-1.46142
Local bank adoption on non-local adoption: Method 1	Constant	.256986	.376214	1.25731
	Non-local adoption	.214177	.0798463	-1.44435
	R^2	.0850278	.0807256	-.316942
Local bank adoption on non-local adoption: Method 2	Constant	.259075	.394988	1.44929
	Non-local adoption	.221951	.0923348	-1.38595
	R^2	.0921889	.0813613	-.631217

Table 6: Economic Significance of Parameters

Change	% of banks adopting	% of consumers adopting	% of transactions completed with ACH
Estimates	68.4%	17.8%	16.2%
No mean bank fixed costs	94.8%	17.9%	17.0%
No mean consumer fixed costs	96.8%	55.5%	52.8%
Always in bad equilibrium	68.4%	17.8%	16.2%
Always in good equilibrium	69.0%	17.8%	16.2%
Local banks internalize externality	83.4%	17.9%	16.9%
All consumers internalize externality	68.4%	17.8%	16.2%

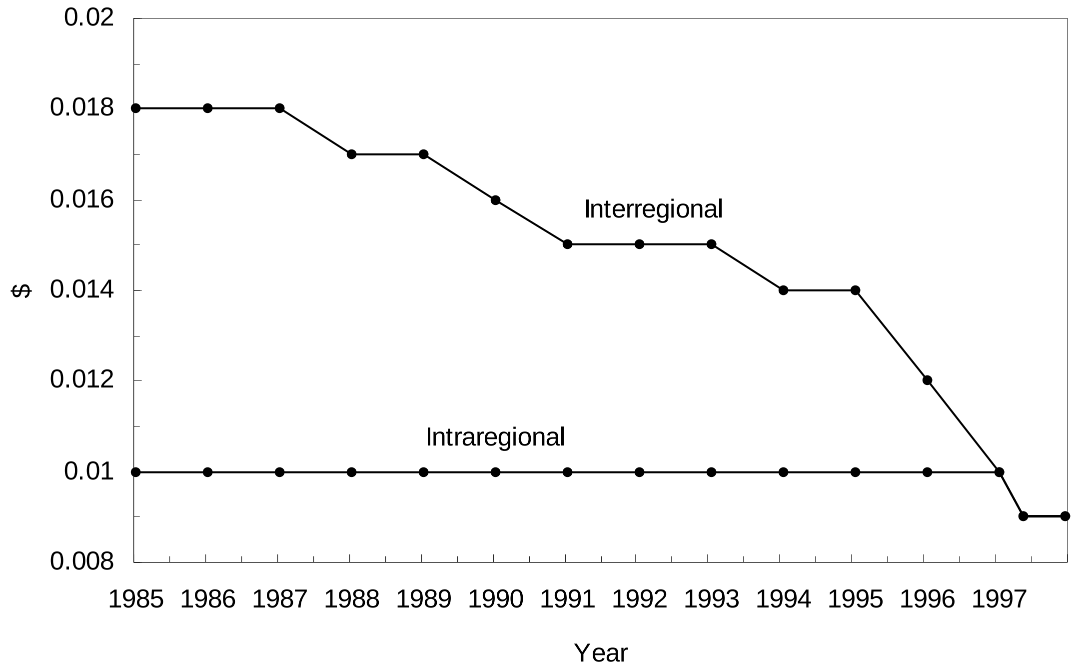
Table 7: Policy Experiments

Policy	None	Subsidize Consumer Mean FC	Subsidize Bank Mean FC	Very Large Consumer Subsidy	Very Large Bank Subsidy	Subsidize 0.17 Cons. Mean FC	Large Bank Subsidy + 0.31 Cons FC subsidy
% of Local Banks Adopting	68.4	96.8	94.8	99.9	100	76.3	100
% of Consumers Adopting	17.8	55.5	17.9	93.8	18.0	22.9	27.9
% ACH transactions	16.2	52.8	17.0	89.6	17.1	21.2	26.7
Firm Profits	1.36 million	4.67 million	1.48 million	8.00 million	9e10	1.81 million	9e10
Consumer Utility	36921	90831	38917	31 million	39251	42620	49222
Cost to Govt. (in profit units)	0	0	53513	0	9e10	0	9e10
Cost to Govt. (in utils)	0	80585	0	4.46e11	0	5669	12564
Total Profits	1.36 million	4.67 million	1.43 million	8.00 million	1.44 million	1.81 million	2.30 million
Total Utility	36921	10245	38917	-4.46e11	39251	36951	36921



Maine
COLEBROOK
MORRISVILLE
Vermont
LEBANON
New Hampshire
AUBURN
CAMDEN
DAMARISCOTTA
Manchester
Nashua
Lowell
Lynn
Newton
Quincy
Boston
Massachusetts
Springfield
Brockton
PROVINCETOWN
Providence
Hartford
Rhode Island
Connecticut
New Haven
NANTUCKET
Stamford
Norwalk

Figure 2: Per-item origination fees for Federal Reserve ACH Processing



Note: In May 1997, volume-based pricing was introduced, with price set to 0.9 cents per item for files with less than 2500 items and 0.7 cents per item for files with 2500 or more items.