Coordination vs. Differentiation in a Standards War: 56K Modems*

Angelique Augereau
McKinsey and Co.

Shane Greenstein
Kellogg School of Management

Marc Rysman
Boston University

November 27, 2004

Abstract

56K modems were introduced under two competing incompatible standards. We show the importance of competition between Internet Service Providers in the adoption process. We show that ISPs were less likely to adopt the technology that more competitors adopted. This result is particularly striking given that industry participants expected coordination on one standard or the other. We speculate about the role of ISP differentiation in preventing the market from achieving standardization until a standard setting organization intervened. JEL: L15, L63, L86

1 Introduction

This paper studies the adoption of 56K modems by Internet Service Providers. Introduced in 1997, 56K modems allowed for data transfer off of the Internet at up to twice the speed of the previous technology at a time when the demand for large files such as graphics became increasingly important. Originally, there were two competing specifications for the standard from two competing consortia, one led by equipment manufacturer, US Robotics, the other by Rockwell. The

*We thank seminar audiences at the SED 2002 meetings, the University of Toronto, Brown University, the Department of Justice, the University of Illinois and Pennsylvania State University, the AEA 2003 meetings, the Federal Reserve Bank of Chicago, Northwestern University, Syracuse University and the Graduate School of Business at Stanford University. We also thank Tim Bresnahan, David Dranove, Phil Haile, Mike Mazzeo, Ariel Pakes, Greg Rosston, Katja Seim, two anonymous referees and the editor, Joe Harrington, for helpful comments. Martino De Stefano provided excellent research assistance. This research was supported by NSF Grant SES-0112527 and a grant from NET Institute.

technologies were functionally identical in the sense that they had the same performance characteristics. However, these technologies were incompatible. If a consumer used one technology and the consumer's Internet Service Provider (ISP) used the other technology, then data transfer speed diminished to that of the previous technology, only 33K or 28K.

We focus on understanding the role of competition in adoption by an ISP. We show that ISPs differentiated across technologies rather than coordinating on one technology or the other. Specifically, we show that ISPs were less likely to adopt a technology as more of their local competitors adopted that technology. This differentiation is particularly important because, as we discuss below, it hindered coordination on a single standard technology, which could have provided important benefits.

Theories about standardization discuss the role of competitive choice between standards, but few prominent cases ever permit researchers to garner a close look at behavior during deployment, as we get here. Related, very little empirical research examines how competitive behavior shapes demand for competing standards or *vice versa*. Data needs are the primary impediment. We rarely observe competition between two comparable technologies played out in more than one market. Even when that occurs, it is often difficult to disentangle the effects of competition from other important effects.

This study's setting is uniquely well suited to meet these requirements. An important feature of the ISP market is that consumers almost always connect to ISPs within their local telephone calling plan. This creates numerous geographically distinct or partially overlapping markets, which leads to geographically dispersed decision making and a variety of competitive interactions. We study over 2200 ISPs in 2300 calling areas. Thus, we are able to compare decisions across markets, where a variety of factors shape decision making, such as the competitive and demographic environment and ISP size.

We employ a series of empirical approaches. Simple statistics illustrate a prevalence towards "even splits" in local markets. That is, adopting ISPs were more likely to be evenly split between the two technologies than would be predicted by independent random choice. We also estimate a structurally motivated model of each ISP's choice over adoption of the two technologies as a function of the competitive environment, local demographics, ISP characteristics and ISP decision-making across multiple markets. We capture the influence of an ISP's adoption decision on its rivals in a discrete game of imperfect information as suggested in Seim (2004).

Estimating a discrete game has well-known problems with endogeneity and multiple equilibria. Fortunately, our data contain a great deal of useful variation. ISPs have different technological characteristics and face different demographics characteristics due to imperfectly overlapping coverage areas. We use this exogenous variation to predict the number of competitors that a given ISP faces. We are particularly sensitive to the robustness of our inferences to unobservable errors at ISPs or at locations, so we pursue a variety of strategies for estimation in the presence of such errors. The great asymmetry in the data means that our models typically have a unique equilibrium. Throughout we conclude that an ISP is less likely to adopt a technology as more its competitors do so.

Understanding the deployment 56K modems is also interesting because the "standards war" for 56K was well publicized. While many contemporary press reports discussed how modem makers competed fiercely for adoption by the earliest choosers, few have substantial data. Citing such accounts, Shapiro and Varian (1999), pp 267-270, feature the case prominently in their discussion of strategic behavior and consortia development prior to deployment, but, again, they do not present any evidence about actual adoption. Similarly, contemporary press accounts tend to cover announcements from firms, not the deployment in each local areas. No research has closely examined the deployment decision of service providers, as we do.

The events are also interesting because they end in intervention from the

International Telecommunications Union, a quasi-government agency. Before ITU intervention, this experience appeared to be an example of "coordination failure." That is, there was a benefit to coordinating ISPs and consumers on a single standard as quickly as possible, but market actors failed to quickly standardize. Market participants expected that standardization would arise because it was in user interest. The popular standard would have more ISPs servicing it, which ensured consumers of high-quality, low hassle, low price service into the future. However, coordination did not arise in the first year of competition. Not only did the two technologies maintain relatively similar levels of ISP commitments, but overall sales to consumers and ISPs were well below what the market could have supported. Sales increased only after the ITU introduced a third incompatible standard as a new focal point. The new standard quickly gained market acceptance and high industry sales followed.

Did competitive rivalry among ISPs contribute to the market's inability to coordinate on a standard unaided? In a final more speculative section of the paper, we argue that the standards war was prolonged by the combination of the market's structure and the behavior it induced, i.e., geographically dispersed decision making and incentives to differentiate locally. This discussion directs further attention at issues not highlighted in the applied literature on standardization, such as the role of service provider competition in a standards war, and the role of standard setting organizations when the specification for the potential standard is not fixed.

2 Related Literature

While user choice between alternatives plays a prominent role in many models of standardization, there are few empirical studies for characterizing its effect. Those that do so focused on either the decision of whether to adopt a standard or not, or the decision of which standard to adopt, but has not dealt with decisions linking choice between two standards and non-adoption.

A few prior studies of competition in technology adoption provide us with general approaches for measuring competitive incentives. For instance, Klepper (2002) uses exit patterns, comparing very competitive and oligopolistic markets, to suggest a strong role for competition in cost-reducing technology adoption in a number of manufacturing industries. Genesove (1999) provides a study of the adoption of offset printing by newspapers and argues that firms in more competitive markets adopted earlier. Mulligan and Llinares (2003) show that ski-lifts were less likely to adopt quality-enhancing technology when local competitors had done so.

We borrow broadly from the general approach of empirical studies of technology adoption in network industries.[1] For example, Saloner and Shepard (1995) show the existence of network effects in bank service by showing that banks with more consumers adopt ATM networks earlier. Like our paper, they infer consumer behavior from observing decisions by firms in different locations. Gowrisankaran and Stavins (2004) and Ackerberg and Gowrisankaran (2003) look at the adoption of automated clearinghouse technology by banks. As with our paper, they exploit overlapping local geographic markets for important variation. However, they use a very different structural model of adoption incentives.

There are a few papers that empirically model horizontal competition between two standards. For example, Dranove and Gandal (2003) argue that the introduction of the DIVX standard slowed down the acceptance of DVD technology. Park (2000) and Ohashi (2003) study the standards war between VHS and Beta in the VCR market. Using market level data on quantities and prices, they focus on the role of installed base. Gandal, Kende and Rob (2000) study network effects between producers of compact disks and producers of CD players.

[1] Shy (2001) provides an introduction and overview of both theoretical and empirical work on network effects.

They too find evidence of the interaction between software and hardware.

There also is a small literature on the behavior of ISPs. Closest to the present study is Augereau and Greenstein (2001), who examine the upgrade behavior by ISPs into higher speed technologies. The statistical model compares the behavior of ISPs located in "urban' and "rural' locations. They find faster upgrade behavior by larger ISPs and those who have experience investing in related technologies. Locating in urban settings plays a role too, but it is less important. Using related data, Greenstein (2000) looks at whether ISPs differentiate from each other by offering services other than dial-up basic access, such as web design, network maintenance, hosting, or broadband to business. In that data ISPs in more urban settings tend to have more advanced services. In both cases it not possible to distinguish the effect of more competition from other factors that vary over the cross-section of locations.

In comparison to the previous literature this paper offers advances in several respects. First, we have access to very detailed data about who competes with whom in ISP markets. This permits a direct measurement between the adoption decision of one competitor and all rivals. Second, we provide an empirical model of technology adoption in a standards war. Third, we provide an empirical method of a discrete decision while controlling for econometric issues of endogeneity, using the insights from the model of Seim (2004) to aggregate over the decisions of many other rivals. As such, the method generalizes to other types of decisions over discrete choices in rivalrous settings.

As noted, Shapiro and Varian (1999) discuss the tactics by sponsors of competing standards, both in general and in the particular case of 56K. Their general discussion focuses on characterizing a standards battle and deducing whether tactics are aligned with market incentives. Their discussion of the 56K modem case is broadly consistent with ours. However, they do not offer a detailed explanation of why there was adoption failure prior to the ITU decision and do not consider incentives to differentiate across the standards by ISPs. Shapiro and Varian also suggest a faster diffusion process then we do prior to the ITU decision, relying mostly on announced adoption decisions for source material. Most of these announcements come from the medium to large firms. As they point out, many of these announcements were not followed by deployment. In this study, we document this activity in the context of all deployment behavior.

Finally, on a broad level we also build on a literature of cases studies of other industries that experienced events that looked like coordination failure to contemporaries. For example, Postrel (1990) associates the failure to adopt quadraphonic stereo with the presence of multiple, competing standards, which created confusion and delay downstream in distribution. Saloner (1989) attributes the failure to unify on a single standard of UNIX in the 1980s to proprietary interests in pursuing strategies that raise switching costs to work station users. Besen and Johnson (1986) and Rohlfs (2001) also relate a number of stories of delayed or failed adoption. For example, Besen and Johnson report that AM stereo required broadcasters and radio owners to be on the same standard, and broadcasts were delayed by the presence of multiple standards. As in our case, these are examples in which increasing the number of new choices plays a role.

3 Industry[2]

A modem allows a computer to send and receive data over a telephone line. Up until early 1997, 33.6K was the fastest modem available for use with analog telephone lines. A 33.6K modem can send and receive 33.6 kilobits of data every second. Most modems connect to the Internet through a local telephone call to an Internet Service Provider. In 1997, about 93% of the U.S. population had access to a commercial ISP (Downes and Greenstein, 2002). As ISPs and telephone companies upgraded their connections to each other, it became technologically possible to raise modem speeds to 56K. The concurrent devel-

[2] A more in-depth discussion of these issues can be found in Rickard (1997a, 1997b, 1998).

opment of the World Wide Web and the use of more graphics increased demand for faster Internet access,providing demand for 56K technology.

Players in the modem industry fell into two camps, either with US Robotics which developed the "X2" modem[3] or with Rockwell Semiconductor which called their product "K56Flex." Both brought their product to market at essentially the same time, February 1997. Independent comparisons showed that the two technologies worked equally well, although there was significant variability across and between technologies depending on local connection characteristics. The two technologies were incompatible in the sense that a consumer with one standard that connected to an ISP with the other standard would receive data at only 28 or 33K (at most).

The ISP market was young in 1997, undergoing growth in new users and new entry in service providers. There were thousands of small firms with very small geographic focus, a few hundred firms with service beyond one city, and a few dozen with national or near-national footprints (Downes and Greenstein, 2002). Only the large firms had recognizable brands, such as AT&T Worldnet or America-On-Line. Small and medium ISPs often offered other Internet services in addition to their dial-up service. Many took strategic positions as early movers into new technology and new services as a way to develop local customers bases and differentiate from their branded national ISP rivals (Augereau and Greenstein, 2001).

The cost of the new modems depended on the purchaser. Modems for consumers were initially priced at around $200, as compared to 33K modems around $100.[4] For ISPs the conversion depended on their technology. The typical installation required a Remote Access Server, a large server that came equipped

[3] "X2" referred to the fact that 56=28X2. Although modems were up to 33K, much of the market was at 28K, and 33K used the same basic technology at 28K.

[4] New PCs accounted for only a fraction of demand for faster modems in one year. Sales to the installed base of existing PC users through retail outlets was potentially much bigger.

with high quality modems and required T1 lines or ISDN lines.[5] Such systems cost more than $50,000 to install 50 ports.[6] Many ISPs had already invested in Remote Access Servers and T1 or ISDN lines for some ports as they were also an efficient way to handle 33K modems. For those ISPs they could simply upgrade their server. Doing so cost $50 to $100 per port and was sometimes offered for free as the standards battle intensified. The ability to upgrade depended on the server – USR servers could be upgraded only to X2, most other servers could be upgraded only to Flex. The result is that upgrade costs were much higher for some ISPs than others, and varied across standards. ISPs often used complicated combinations of servers and consumer-grade modems from multiple vendors, so it is unlikely that rivals knew each other's technology exactly. We use these features to motivate our assumptions about unobserved terms in the structural model.

Before moving forward, we briefly discuss the importance of coordination and differentiation in this industry. Network effects could exist at a locality to the extent that consumers purchasing one modem technology make the technology more attractive to other consumers. This can happen via an indirect effect: consumer purchase attracts more ISPs to service that technology which then attracts more consumers. Consumers prefer a technology to have more ISPs servicing it for a number of reasons. A standard served by more ISPs should have greater price and quality competition.[7] Relatedly, more ISPs on a standard lowers switching costs for consumers and ensures the consumer that they will not be technologically "stranded" if their ISP shuts down. Also, coordinating

[5] T1 and ISDN lines are fast, digital connections to the telephone network.

[6] Each connecting consumer requires 1 port. Allowing for the fact that consumers do not all connect at once meant that ISPs typically required 1 port for every 3 or 4 consumers. The number of ports that a typical ISP maintained at a given point-of-presence ranged from 50 to many thousands.

[7] There is little empirical evidence on this issue. Greenstein (2000) shows that ISPs in more competitive markets provide higher quality service but does not establish the direction of causation.

on a standard reduces the difficulties solving technical problems that interfered with achieving higher speeds, which was salient to both ISPs and end-users who were learning how to deploy this new communications standard.

These reasons differ somewhat from standard models of indirect network effects (for example, Church and Gandal, 1992) in which consumers prefer a large number of service providers in part because they plan on purchasing from all of them. Most likely, modem purchasers are not interested in service provider variety for its own sake. However, the effect of the number of ISPs on pricing and quality is enough to generate a network effect. Even without a preference for variety, it still can be true that consumers attract ISPs and ISPs attract consumers. Rysman (2003) provides an example of this.

Because we do not observe consumer data, we cannot measure the empirical importance of the network effect to users. However, it is clear that network effects must have been important. As detailed in Greenstein and Rysman (2004), all written and oral accounts of the industry that we collected indicate that all knowledgeable market participants were sure that the market would eventually coordinate on a single standard and that consumer demand would grow substantially in that event. Subsequent events proved these predictions correct, suggesting that network effects were strong.

Conversely, ISPs prefer differentiation for some of the same reasons that consumers prefer coordination. When switching costs lock consumers into a particular modem technology at least over some time span, it can be beneficial to serve a market with few competitors even if doing so implies serving a smaller potential market of consumers. Fewer competitors should imply less price and quality competition and higher margins. Even if there is a chance that the market will tip towards the more popular product in the future, the long-run gains of serving the large market may be outweighed by the short-run benefit of serving a less competitive market.

In what follows, we interpret an even split across technologies between ISPs as evidence of competition between ISPs. An alternative explanation may be that there were endogenous responses by technology sponsors. That is, we may observe ISPs split across technologies because a sponsor of a technology may lower prices in a location where it is failing. However we see no evidence that this occurred in the contemporary press and, as resale of equipment does not seem difficult, it may be hard to implement a strategy of geographic price discrimination.

4 Data

The data set used in this paper draws on a number of sources. The unit of analysis is the ISP and we use two directories of ISPs to create our data set. The first is from *theDirectory*. The list from *theDirectory* is meant to be comprehensive, including even the smallest ISPs. Importantly, *theDirectory* provides each phone number that each ISP can be contacted through, so we are able to determine each ISP's points of presence (POPs). However, *theDirectory* does not provide any other data on ISPs. In contrast, the *Boardwatch* directory gives information about the technologies that ISPs were using – in particular, which type of 56k an ISP adopted in October 1997 and July 1997. These dates are quite early in deployment of new modems, before many large ISPs adopted and before it was apparent when the ITU would produce a standard. Also, *Boardwatch* lists whether an ISP had a T1 line and whether an ISP offered ISDN service to consumers. However, *Boardwatch* does not provide information on individual POPs. We merge the two data sets so we have both ISP technologies and their geographic locations.

This merge has a number of implications. First, we lose many observations from *theDirectory* because *Boardwatch* is less comprehensive. However, we believe that this loss is not a serious problem as *Boardwatch* contains data on the "most important" ISPs. ISPs that are not in *Boardwatch* were unlikely to adopt

56k. We assume the 56k adopters face a "competitive fringe" of 28/33K firms to which these "lost firms" belong.

A second implication of the construction of our data set is that we observe only one adoption decision for each firm. We do not see if an ISP adopted one type of 56k in one market, the other in a second market, and chose not to adopt in a third market. Again, we believe this issue is not problematic as it appears that ISPs themselves treated the adoption decision as a single firm-wide decision. There are a few reasons for this to be the case. ISPs had an incentive to deliver uniform service throughout their market area, especially for clients who traveled. The choice of standard even seemed to become part of the marketing campaign of some ISPs. Also, competition among Rockwell and US Robotics led to the offering of exclusive contracts to ISPs. For instance, it is clear from press releases that national ISPs such as AT&T and AOL, when they finally did adopt in November 1997, adopted only a single technology and did so throughout their service areas. Shapiro and Varian (1999) also note the coalitions that the technology sponsors tried to form across ISPs and equipment providers.

Matching the October releases for each data set gives us 2233 ISPs.[8] Next we determine markets for ISPs. Consumers almost always work with an ISP that is within the consumer's local telephone calling range. From CCMI, we obtain the *Qtel* data base which allows us to link telephone numbers to telephone switches, and switches to local calling plans. We assign each switch to the primary consumer local calling plan available from the incumbent telephone company. From this information, we can determine the switches that are served by each ISP, and the competitors that a consumer at each switch could potentially call. Also, we observe the zip code associated with each switch, which we use to add demographic data. We match switches to zip codes and counties and

[8] The original *Boardwatch* data set had 2653 observations, whereas theoriginal list from *theDirectory* contained 5363 ISPs.

use zip code level demographics from the 2000 Decennial Census or (when zip code level data was unavailable) county level demographic data from the 1995 USA Counties CD-ROM. ISPs are spread over 9,076 switches, which creates 216,583 separate ISP-switch combinations.[9]

Using switches as a measure of size shows that the Internet access market is served by many small ISPs and a few very large ones. The mean number of switches served by an ISP is 96.8 but the distribution around the mean is very skewed. The median ISP serves 16 switches (1 or 2 local calling areas), the 75th percentile is 32, and the largest 5 firms serve more than 4000 switches each. Note that there are more than 9,000 switches in the data set so no ISP covers the entire market.[10]

Table 1 shows adoption rates in July and October. Note that we construct the data using the October samples and simply append the ISPs' choices from July. That is, we ignore entry and exit over this 3 month period. Adoption by July was very low. While there was significant adoption by October, still only about half of ISPs had adopted. Moreover, the vast majority of non-adopting ISPs were large, so the percentage of customers served by 56k was much lower than a half. The slight lead enjoyed by X2 in July had turned into a slight lag by October. While very few firms adopt both technologies in July, they represent more than 15% of adopters in October. Note that having a T1 line is highly correlated with adoption. Among ISPs with T1 lines, 56% adopted X2 or Flex, whereas the adoption rate is 38% for those without a T1 line. We

[9] Note the improvement over Augereau and Greenstein (2001), who also examines ISDN and 56k modem adoption. That paper uses county boundaries as the geographic market, which identifies urban and rural locations, but not direct competitors. In contrast here, the emphasis on the greater detail identifies direct rivals among potential adopters.

[10] There is no such thing as a truly national ISP in the United States at this time. Most switches are served by ISPs, but by ISPs that we classify as being unlikely to adopt 56k: Downes and Greenstein (2002) show that the largest ISPs are present mostly in urban areas. Thousands of switches in rural areas have only one of two local ISPs, if they have any.

	July 1997		October 1997	
None	1909	85.5%	1136	50.9%
X2	185	8.3	389	17.4
Flex	112	5.0	523	23.4
Both	27	1.2	185	8.3

Table 1: Number and Percent of ISPs Adopting

also observe whether firms offer ISDN lines to their consumers.[11] Firms offering ISDN service adopted 56K 66% of the time, whereas the adoption rate was only 29% among firms that did not offer ISDN service.

A simple way to look at the data is to take local calling areas as distinct markets. Doing so is complicated by the fact that local calling areas do not create a partition of the United States – there are areas where switch A can make a local call to switch B and switch B can make a local call to C but A and C are not in the same local calling area. Hence, we create local calling areas by making some arbitrary assignments of switches to calling areas when a question arises. We find that this arbitrariness is not very problematic for looking at some simple summary statistics.[12] Moreover, our final estimation procedure properly accounts for the overlap patterns.

Our method creates 2,298 local calling areas. Local calling areas have relatively few firms in each one. The average number of ISPs in a calling area is 15 with a standard deviation of 20.8. However, there are 738 calling areas with only 1 ISP and the median number is only 3. Table 2 gives average adoption rates by local calling area. Again, there are only a few adopters in each calling

[11] Our dummy variable indicates if a firm offered ISDN service to consumers, not if a firm has an ISDN connection to the Internet. Firms that offer ISDN service to consumers require an ISDN connection to the Internet themselves. But many ISPs had ISDN connections without offering ISDN service to consumers.

[12] Most of the issues arise in dense urban markets with many competitors. Medium to low density locations make up the bulk of the dataset and these are not problematic.

	7/97	10/97
ISPs	15.06	15.06
Adopters	0.99	5.98
X2	0.59	2.58
Flex	0.22	1.99
Both	0.18	1.40

Table 2: Averages by Local Calling Area

area. The average number of adopters in October 1997 is about 6. Interestingly, although Flex leads X2 when tallied by ISP (as in Table 1), X2 leads Flex when tallied by locale (as in Table 2).[13]

5 Simple Measures of Differentiation

Our goal is to show that ISPs differentiated across technologies instead of coordinating on one technology or the other. In this section, we present simple statistics that capture the spread of ISPs across the two technologies within local calling areas. This "first cut" of the data shows that ISP adoption is characterized by differentiation. This finding is also reaffirmed in the more comprehensive empirical model presented in Section 6.

Our approach in this section is to compare the national adoption rate with the adoption rate in each local calling area. If the rates are close to the same, it suggests that ISPs were differentiating from each other. If local markets are characterized by agglomeration on one standard or the other, it suggests network effects were important.

Let $\{n_i^0, n_i^A, n_i^B, n_i^{AB}\}$ be the number of ISPs in market i (a local calling

[13] Although this finding suggests that larger firms were more likely to adopt X2, we find the parameter on size (POPs) to be very similar across the two standards in our final estimation procedure.

area) that do not adopt, that adopt Flex, that adopt X2 and that adopt both. We start with a graphical presentation of our data. We calculate the number of adopters of only X2 as a percentage of the number of ISPs that adopt only one technology, ignoring markets with only one such firm. That is, we compute $n_i^B/(n_i^A + n_i^B)$ in each market where $n_i^A + n_i^B > 1$. For now, we ignore firms that adopt both or neither technology (n_i^0 and n_i^{AB}). The national adoption rate computed in this way is 58%.

Figure 1 presents a histogram of adoption by calling area and captures the unconditional patterns of greatest interest for this study. The black bars represent the observed adoption rates (aggregating over markets of different size). Figure 1 shows that most of the calling areas have between 50% and 80% adoption rates of X2, and there are very few calling areas with adoption close to 0 or 100%. As a point of comparison, we also calculate what would have happened if ISPs made independent random choices – that is, if $n_i^A + n_i^B$ firms in each market chose between A and B independently with probability 58%. These results are represented by the gray bars. Figure 1 shows that independent random choice puts less weight in the center of the distribution and more weight on the tails. The black bars are higher than grey bars for the middle three bins and are lower than the grey bars for the outer seven. Figure 1 suggests that differentiation (even splits between each technology within each locale) characterizes this data, relative to independent random choice or coordination.

In order to test statistically whether the hypothesis of independent random choice can be rejected, we develop a statistic we call the Multinomial Test of Agglomeration and Dispersion (MTAD), formally presented in Rysman and Greenstein (2004). In this circumstance, the test calls for computing the likelihood of observing the vector of n_i^A "successes" in $n_i^A + n_i^B$ "trials" as generated by a binomial distribution with success probability $p = 0.58$. We compare the likelihood value to what we would have gotten if n_i^A was actually drawn from a set of independent random trials. Agglomerated outcomes leads to a lower

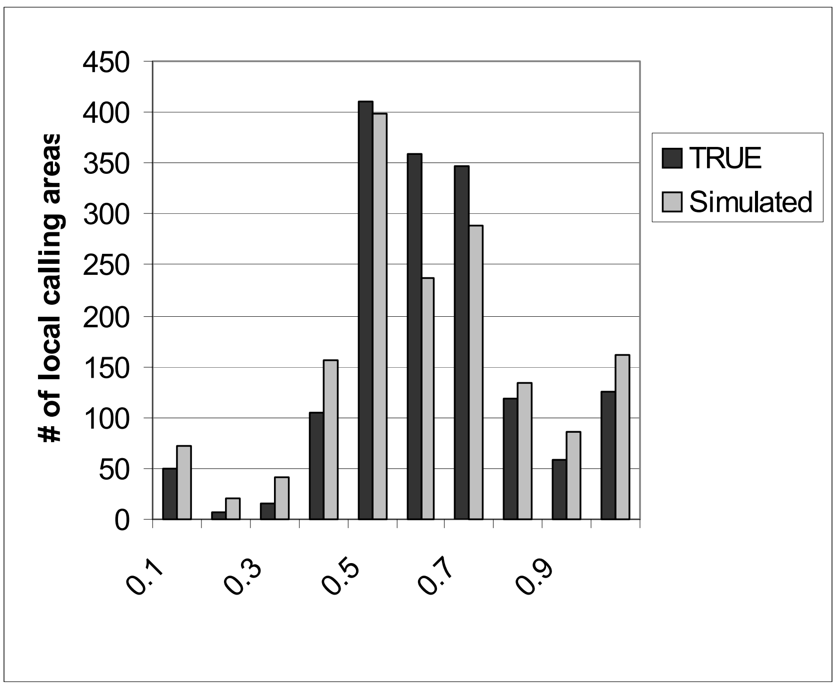

Figure 1: Percentage of ISPs adopting X2.

likelihood value than would have been generated by independent random choice. Dispersed outcomes leads to a higher likelihood value. To see why this is the case, consider the combinatoric expression $\binom{t}{s}$ in the binomial likelihood. For this example, suppose the number of trials $t = 4$. If s is the number of "successes" drawn with independent probability $p = 0.58$, then $E[\binom{4}{s}]$=2.9. If the outcomes are agglomerated ($s = 0$ or $s = 4$) then $\binom{4}{s} = 1$. If the outcomes are dispersed, we have $\binom{4}{2} = 6$. In this circumstance, MTAD produces the same results as the dartboard index of Ellison and Glaeser (1997). [14] The appendix presents functional forms for MTAD.

Results appear in Table 3. In Row 1, we consider ISPs that adopt only one technology and markets with at least two such ISPs. We report the log-likelihood of the observed data arising from a binomial distribution averaged over markets. We also report the expected log-likelihood and standard deviation if the outcomes really were generated by a binomial distribution. The observed log likelihood is significantly higher. That supports the result in Figure 1 that the data is characterized by differentiation rather than by independent choice or agglomeration. Row 2 adds ISPs that adopt both technologies to the "X2" group and Row 3 adds those ISPs to the Flex group. The result is unchanged.

Row 4 divides firms into those that adopt any technology and those that do not adopt. Again, we see that differentiation describes these outcomes. We might see the opposite result if demographic data was important for adoption as some areas would have high adoption and others would not. This result foreshadows our finding below in the more comprehensive model where we find that demographics play a limited role. It is possible to consider more than two outcomes with MTAD. For instance, one could check if ISPs differentiated across the four outcomes {0,A,B,AB}. Rysman and Greenstein (2004) present

[14] The dartboard index is based on a structural model of agglomeration that cannot naturally be applied to dispersed outcomes, whereas MTAD handles agglomeration and dispersion. MTAD is discussed in greater detail in Rysman and Greenstein (2004).

Table 3: Dartboard Test for Differentiation

Description	Observed Likelihood	Expected Likelihood	Std. Dev.	Markets
Adopt only X2 vs. Adopt only Flex	-1.460	-1.631	0.016	1595
Adopt X2 or Both vs. Adopt only Flex	-1.541	1.660	0.015	1698
Adopt only X2 vs. Adopt Flex or Both	-1.595	-1.731	0.015	1698
Adopt vs. Not Adopt	-1.825	-1.912	0.013	2200

these results and they find differentiation again, as would be expected based on Table 3.

Strikingly, outcomes appear differentiated regardless of market structure. Whereas one might imagine different forces at work for small and large markets, we find differentiation for all market sizes. We compute MTAD as a function of the number of ISPs in each market separately. Figure 2 plots the difference between the observed and expected log-likelihood, normalized by it's standard deviation by number of firms. A result greater than 1.96 can be interpreted as statistically significant. Although the result is not always statistically significant when studying the data at this level of detail, particularly in larger markets, the important result is that the statistic is always above zero, implying that differentiation is prevalent for all market structures.

Another way to look at the issue of differentiation across technologies is to exploit the dynamic aspect of the data. We can compare choices made up to July 1997 to choices made afterwards. The results appear in Table 4. This table shows that there were 1029 local calling areas where there was at least one adopter by July. The columns refer to whether or not X2 was leading in that calling area by July 1997. On the row is the number of calling areas where more ISPs adopted X2 than Flex in the July-October window. For instance, the table shows that of the 686 calling areas where X2 led in July, Flex tied or led X2 over the next 3 months in more than half the calling areas. The numbers are

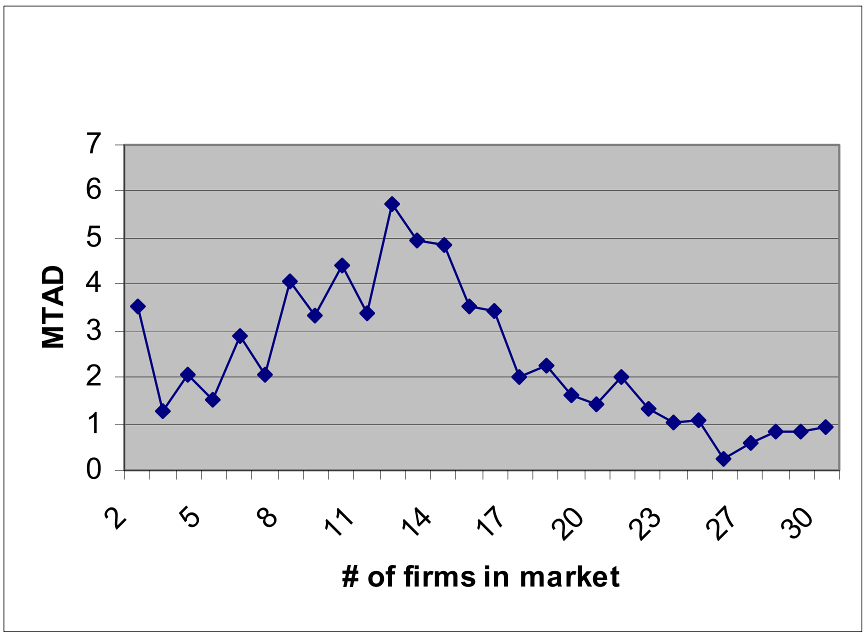

Figure 2: Multinomial test of agglomeration and dispersion (MTAD) by market structure

Table 4: Adoption in July 1997 versus adoption over the next 3 months

		Adoption by 7/97			
		X2>Flex	Tied	Flex>X2	Total
Adoption	X2 Leads for 7-10	325	79	67	471
Betw. 7/97	Tie	178	54	65	297
and 10/97	Flex Leads	183	58	20	261
	Total Calling Areas	686	191	152	1029

more striking for calling areas where Flex led in July. Of these 152 calling areas, there are 3 times as many locales where X2 led for July - October as there are those where Flex led. These numbers suggest that ISPs did not adopt the same technology as the technology that had obtained a lead in their local market.

Overall, these statistics characterized the data by differentiation, not coordination. They are strongly suggestive of the results we find in the full empirical model in Section 6.

6 Estimation

So far the methodology ignores important features of the industry. In particular, ISPs make a single adoption choice at all of their points-of-presence so local calling areas are not independent markets. Furthermore, the methodology above does not exploit demographic data or heterogeneous features of ISPs and is difficult to interpret when we recognize that some firms adopt both technologies. The main econometric model addresses all of these features.

6.1 Model

In our econometric model, ISPs that offer 33K service at a switch decide whether or not to offer 56k service on X2, Flex, both or neither. In this sense, the model is like an entry game into two markets, X2 and Flex, in which we observe potential entrants, as in Berry (1992). We model the entry game as one of imperfect information, where we allow for ISPs to observe their own unobservable draws but not those of their competitors. In this regard, we follow the estimation methodology of Seim (2004).

We use a bivariate probit model specialized to control endogenous variables. Our specific estimation model is as follows. There are N ISPs and I locations. Locations in the model are equivalent to switches in the data.[15] The set of locations in which ISP j appears is ϑ_j. We compute ϑ_j to be the set of all switches from which ISP j can be contacted by a local telephone call. An ISP may adopt either standard s, adopt both or not adopt, but the ISP makes the same adoption decision in every location. The number of ISPs that have adopted technology s at location i besides j is n_i^s, where s can be equal to $\{0, A, B, AB\}$ (none, Flex, X2, and both) and j will be obvious in context. The potential adopters at a switch are identified by finding all ISPs that can be contacted by a local telephone call.

ISPs draw a cost shock ε_j^s for each technology. ISPs observe their own draws of ε_j^s but not those of their competitors. The shocks represent the adoption cost for the ISP. One source of these costs may be the combination of servers, consumer grade modems and digital connections an ISP has that, as argued earlier, affected the adoption cost of 56K and was arguably not well known to competitors. Because of this incomplete information, we search for a Perfect Bayesian Equilibrium.[16] Let the vectors x_i^l and x_j^f capture location specific and ISP specific variables. The expected profit from adopting A or B in location i

[15] In this section, location i refers to a switch whereas in Section 5, location i referred to a local calling area.

[16] Our solution concept is equivalent to a Quantal Response Equilibrium. See McKelvey and Palfrey (1995), and Haile, Hortascu and Kosenock (2003) for a discussion of its empirical content.

for ISP j is:

$$E[\pi_{ij}^s] = x_i^l\beta_1 + x_j^f\beta_2 + E[\psi_1(n_i^s + 1) + \psi_2 n_i^{AB} + \psi_3 n_i^{-s}|\mathbf{X}, \theta] \quad s = A, B \quad (1)$$

The expectation term captures the effect of competition. The parameters ψ_i measure the effect of competition on profits. The "+1" in the parenthesis accounts for the effect of ISP j on profitability in location i. The matrix $\mathbf{X}$ contains all exogenous variables, including all values of x_i^l, x_j^f and adoption decisions by ISPs in previous periods. Average profit for ISP j from adopting technology s is:[17]

$$E[\Pi_j^s] = \frac{1}{I}\sum_{i\in\vartheta_j} E[\pi_{ij}^s] + \varepsilon_j^s, \; s = A, B$$

The variable ε_j^s is a random fixed cost for technology s unobserved by the researcher or other ISPs. We assume that $[\varepsilon_j^A \; \varepsilon_j^B]$ is distributed *iid* according the standard bivariate normal distribution with correlation parameter ρ. We assume that economic profit in the 33K market is constant across markets and normalize this profit level to zero. We add a term for economies of scope when an ISP adopts both technologies. That is:

$$E[\Pi_j^0] = 0 \;\; E[\Pi_j^{AB}] = E[\Pi_j^A] + E[\Pi_j^B] + \delta$$

The parameter δ represents the additional payoff of adopting both technologies beyond the sum of adopting each technology and may be negative. The ISP chooses the option with the highest expected payoff. The parameters $\theta = \{\beta_1, \beta_2, \psi_1, \psi_2, \rho, \delta\}$ are to be estimated. In practice, we can allow them to differ for each technology. Our goal is to check whether $\psi_1 < 0$, which implies that ISPs prefer differentiation to agglomeration.

Distinguishing between economies of scope (δ) and a correlation in error (ρ) is difficult. Manski (1998) shows that an "endogenous effect" and a "correlated

[17] Using average profit is equivalent to using total profit and imposing a particular form of heteroskedasticity.

effect" cannot be distinguished in a linear model and terms this the "reflection problem." While these parameters might be identified in our non-linear model, we do not claim that our data distinguish between these effects. In practice, we make the restriction $\delta = 0$ and interpret ρ to capture both correlation in error and economies of scope.

We do not attempt to characterize or solve the full dynamic game here. Instead, our model is static and simultaneous. However, one can think of this paper as estimating a single period of the wider dynamic adoption game. In the single period, there are two stages: the first in which ISPs realize ε_j^s and a second in which ISPs make adoption choices. In describing profits, we have left off any terms representing value in future periods. In this sense, the parameters we estimate capture both the flow profit from their associated variable and the expected future profit. In addition, they may capture the fact that waiting has different values to different ISPs depending on exogenous variables. Capturing dynamic industries with static models is standard in the empirical entry literature. A strength of our paper is that we model decisions made over a 3-month period, a period short enough that a simultaneous game might be a reasonable model.[18]

A widely recognized problem in the empirical literature on entry games is the potentially endogenous determination of the number of competitors. A high draw of ε_j^s might be due to the fact the location is desirable. In that case, n_i^s will also be high. Conversely, a high draw of ε_j^s might suggest that n_i^s is small because ISP j will almost surely enter so that in the equilibrium of an entry game, competitors may not enter. However, as shown in Seim (2004), modelling the entry game as one of imperfect information addresses this issue. In this case,

[18] Mazzeo (2000) considers a case where there is no uncertainty and firms first choose entry and then choose type. We model only the choice of type and take entry as exogenous, which can be thought of as the second stage of Mazzeo's model in the presence of risk-neutral expectations about the choices of rivals.

ISPs make their decision based not on n_i^s but on $E[n_i^s|\mathbf{X},\boldsymbol{\theta}]$,which depends only on exogenous variables. We discuss computation of $E[n_i^s|\mathbf{X},\boldsymbol{\theta}]$ below.

Integrating over ε_j^A and ε_j^B, the implied adoption probabilities for ISPs are:

$$\begin{aligned} P_{jN} &= \text{Prob}(d_j = N) = \Phi(-\Pi_j^A, -\Pi_j^B, \rho) \quad (2) \\ P_{js} &= \text{Prob}(d_j = s) = \Phi(\Pi_j^s, -\Pi_j^{-s}, -\rho) \quad s = A, B \\ P_{jAB} &= \text{Prob}(d_j = AB) = \Phi(\Pi_j^A, \Pi_j^B, \rho) \end{aligned}$$

Here, d_j is the decision of ISP j and $\Phi()$ is the bivariate normal CDF with correlation parameter ρ. In some specifications, we treat all decisions as symmetric and simultaneous regardless of whether we observe them in the July or October data. However, in some specifications we treat decisions made by July as exogenous data. In these specifications, an ISP that adopted one technology in July has the probability of adopting the other in October defined by:[19]

$$P_{js} = \text{Prob}(E[\Pi_j^s] > -\varepsilon_j^s)$$

The likelihood function for observing the N decisions $d_1, ..., d_N$ is:

$$L(d_1, ..., d_n, \mathbf{X},\theta) = \prod_{j=1}^{N} P_{jd_j} \quad (3)$$

Our technique for computing $E[n_i^s|\mathbf{X},\boldsymbol{\theta}]$ is as follows: For the case in which July adoption is treated as exogenous, the expected number of ISPs making choice s at location i can be divided into the the number that chose s previously (by July 1997) $n_i^{s,pr}$ and the number choosing currently $n_i^{s,cu}$, the expectation of which depends on the adoption probabilities. Let $\Upsilon(i)$ be the set of ISPs present in location i:

$$E[n_i^s|\mathbf{X},\boldsymbol{\theta}] = n_i^{s,pr} + E[n_i^{s,cu}|\mathbf{X},\theta] = n_i^{s,pr} + \sum_{k\in\Upsilon(i)} P_{ks} \quad (4)$$

[19] If we assumed that ε_i^s was perfectly persistent over time, then observing previous adoption affects ones inference on this computation according to ρ. We experimented with accounting for this issue and found that it did not affect our estimates.

In order to calculate the term $E[\psi_1(n_i^s+1)+\psi_2 n_i^{AB}+\psi_3 n_i^{-s}|\mathbf{X},\theta]$, we follow Seim (2004) and exploit the fact that the system of equations 1, 2 and 4 form a fixed point equation, solved by the $N \times 4$ matrix of adoption probabilities $\mathbf{P}$, with element P_{js}. For any given set of parameters θ, the first step in solving for $E[\psi_1 n_i^s + \psi_2 n_i^{AB} + \psi_3 n_i^{-s}|\mathbf{X},\theta]$ is to compute $E[\Pi_j^s]^0$, the value of $E[\Pi_j^s]$ assuming no ISPs adopt. Doing so gives probabilities of adoption $\mathbf{P}^0$ that we use to create an initial guess, $E[\psi_1(n_i^s+1)+\psi_2 n_i^{AB}+\psi_3 n_i^{-s}]^0$. Using this value, we calculate $E[\Pi_j^s]^1$, which generates a new set of adoption probabilities $\mathbf{P}^1$ and a new value $E[\psi_1(n_i^s+1)+\psi_2 n_i^{AB}+\psi_3 n_i^{-s}]^1$. We continue iterating in this way as described by Seim (2004) until there is convergence, thereby finding a fixed point in adoption probabilities for the system of best-response functions.[20] The appendix discusses the computation of the fixed point in greater detail. When we do not assume that previous adoption is exogenous, we set $n_i^{s,pr} = 0$ for each i and s and solve for $\mathbf{P}$ for all ISPs and choice possibilities.

A weakness of our approach is that it does not guarantee that there is a unique equilibrium. There may be multiple matrices $\mathbf{P}$ that solve the system of equations above. In this model, this can occur for two reasons. The first is associated with entry models and can occur if $\psi_i < 0$. Intuitively, there could be an equilibrium where ISP 1 is expected to enter with high probability and ISP 2 with low, and another equilibrium where the opposite is true. The second is associated with network effects and can occur if $\psi_i > 0$. In this case, there can be an equilibrium where most ISPs adopt and one where few ISPs adopt. Our methodology requires either a unique equilibrium or an equilibrium selection mechanism. Our estimation relies on asymmetry in the data to find uniqueness. For instance, while one might expect that there are multiple equilibria between two symmetric ISPs, it would be less likely between a very large ISP and a very small ISP or with two ISPs that partially overlap but serve different competitive

[20] The adoption probabilities P_{jN}, P_{jA}, P_{jB}, and P_{j2} can be defined by only 2 cut-offs, so in fact, the solution to our fixed point system is an $N \times 2$ matrix.

situations in their non-overlapping areas. Our strategy is to estimate as if there was a unique equilibrium and check for multiple equilibria at the resulting parameters. In practice, we always find only one equilibrium (by computing $\mathbf{P}$ starting from many different values), which is not surprising given the great asymmetry in our data.

6.2 Identification

This methodology creates a variable $E[\psi_1(n_i^s+1)+\psi_2 n_i^{AB}+\psi_3 n_i^{-s}|\mathbf{X},\theta]$ that nicely captures our intuition for identification. As with standard instrumenting techniques, it is important that there are variables that affect the expected number of adopters that j faces but that do not otherwise affect the decision of ISP j. In this sense the overlapping calling areas of different ISPs acts as a virtue. Demographics faced by competitors of ISP j but not by j itself provide exogenous variation in the predicted number of competitors that j faces. The relationship between ISPs that provides the best identifying power would be an ISP that is completely overlapped by another ISP, where the overlapping ISP is also present in many other locations. Then, the variable $E[\psi_1(n_i^s+1)+\psi_2 n_i^{AB}+\psi_3 n_i^{-s}|\mathbf{X},\theta]$ for the small ISP depends on a large amount of demographic data that do not otherwise appear in their decision. Cases where two ISPs completely overlap but do not appear in many other locations should not provide strong identification as the same demographic data that affect $E[\psi_1(n_i^s+1)+\psi_2 n_i^{AB}+\psi_3 n_i^{-s}|\mathbf{X},\theta]$ also appear directly in the ISP's' decisions. Similarly, ISPs that appear in many different locations but barely overlap with each other should not provide much identification as these ISPs barely affect $E[\psi_1(n_i^s+1)+\psi_2 n_i^{AB}+\psi_3 n_i^{-s}|\mathbf{X},\theta]$ for each other. Further important exogenous variation comes from the characteristics of ISPs, such as their size and the presence of a digital connection.

Other papers, such as Gowrisankaran and Stavins (2004), have used the intuition that the decisions of geographically large firms can be thought of as exogenous to the decisions of small firms. Our intuition is similar, although we capture exogeneity in a more continuous way – the decision of a large ISP may be exogenous to a medium-sized ISP and the decision of large and medium ISPs may be exogenous to that of a small ISP. With our methodology, we do not have to make an *a priori* decision about which ISPs are exogenous to which. A drawback is that we do not use the actual decision of large ISPs as an exogenous variable, only the prediction of those decisions based on explanatory variables.

The above mentioned forms of exogenous variation differ across ISPs but not across technologies. Without further variation in data, the model predicts each ISP has the same probability of adopting technology A and technology B. Thus, $E[n_i^s + 1]$ and $E[n_i^{-s}]$ are perfectly collinear with the constant term. We address this issue in two ways.

The first is to assume that $\psi_3 = 0$. Holding consumer adoption of each technology fixed, the number of ISPs on one technology should not affect the ISPs on the other technology so n_i^{-s} should be important only to the extent that it captures consumer adoption. That is, holding consumer adoption fixed, more ISPs on one technology does not make it more or less likely that an ISP will adopt the other technology. But there are two effects via consumer adoption. First, more ISPs on a technology may indicate that more consumers have already adopted that technology. Second, more ISPs on a technology may attract consumers (both new purchasers and those who use the other technology) in the future as a result of greater competition. However, these effects may be relatively small in a market where consumer adoption and switching was limited and difficult to predict.

A second approach to identifying ψ_3 is to assume that entry decisions in July 1997 are exogenous and exploit them as an instrument that differs across technologies in the same location. This approach introduces additional considerations. Above, we stated that there are two types of endogeneity that we are concerned about. The first is unobserved location effects and the second is unobserved ISP effects. This instrument is useful to the extent that early adoption reflects ISP effects as opposed to location effects.

If location effects are important and persistent, this instrument would be invalid because previous entry decisions would be correlated with current unobservable factors. However, if location effects affected both technologies equally, we would observe markets grouped into high and low adoption - that is, there would be agglomeration with respect to adoption of any technology (not necessarily the same one). Conversely, if location effects were specific to one technology, we would observe ISPs grouping on one technology and not the other – that is, agglomeration with respect to technologies. In fact, Table 3 shows that we observe differentiation in both respects, not agglomeration. Table 3 implies that location effects are probably not very important for these data.

If unobservable effects are primarily associated with ISPs, not location, then previous adoption decisions by competitors are a useful instrument. Previous adoption can be taken as exogenous to the unobservable shocks realized by any given ISP. Strictly speaking, ISPs that did not adopt in July reveal that they had (possibly persistent) low realizations of profitability. However, in our data adoption is so low in July that this is not very informative and, in particular, it does not point towards one technology or the other. We provide additional analysis below and conclude that this possible "endogeneity" will not provide misleading identification of ψ_3.[21]

[21] Another approach to generating different predictions for adoption probabilities is to allow parameters to differ across technologies. However, we believe that the technologies are functionally identical and that any differences in parameters will be due only to statistical error. We experiment with allowing parameters to differ across technologies as a specification test and do not present it as an approach to identification.

7 Results

As a preliminary step, we begin with an analysis that ignores endogeneity issues. Results appear in Table 5. Relative to Figure 1, Table 5 controls for demographics and ISP characteristics and properly handles both the assignment of switches to local calling plans and that fact that ISPs make a single decision in all locations.

The first column presents a bivariate probit predicting adoption of Flex and X2 as a function only of exogenous variables and with parameters restricted to be the same for both technologies. Results indicate that the most important predictor is the dummy for whether or not we know that an ISP has a digital connection to the Internet. Also, the size of an ISP is important, with size measured by the number of telephone switches to which it can provide service. We find a concave function that has a maximum around 250 switches. This reflects the fact that the smallest ISPs did not adopt, presumably because of scale issues. However, the largest did not adopt either, possibly because they served mass markets comprised of less sophisticated consumer bases that were late adopters of 56K modems. Also, ISPs in switches for which one technology had the lead in July are *less* likely to adopt that technology in October. ISPs at switches with more competitors are more likely to adopt, which we take as evidence of unobserved consumer heterogeneity: There are more ISPs in locations where consumers have high demand for new technologies. Locations with more backbone providers have cheaper access to the Internet for ISPs but we see no effect on adoption. Demographics do not seem very important although they are jointly significant. Adoption is lower in areas with more college educated people and higher in areas with lots of turnover or growth. The parameter ρ is estimated to be negative and significant, implying either diseconomies of scope in adopting both technologies or a negative correlation in unobservable adoption costs.

Results are similar in column II, which includes the decisions of competitors as exogenous regressors. This column is useful for establishing conditional correlations and determining the impact of instrumenting. The first row indicates that technology choice is negatively correlated across ISPs. This result supports the conclusion from Figure 1 that differentiation characterizes these data as opposed to agglomeration.

Surprisingly, row 2 indicates that ISPs are more likely to adopt in situations where their competitors adopt both technologies. As it is difficult to construct a causal explanation for this result, we believe that it reflects unobserved consumer adoption. That is, in markets with high consumer adoption, we see both high ISP adoption overall and a high incidence of ISPs adopting both technologies. This conclusion seems to be in tension with row 4 of Table 3, which we interpreted to say that there was little important observable or unobservable heterogeneity across markets. It seems that the effect of unobservable consumer adoption is not apparent in a simple summary variable such as in Table 3 but is noticeable in this more detailed regression. If unobservable adoption is important, it would raise problems for our second identification strategy as it would mean there may be an important persistent location effect, and that early ISP adoption is not exogenous to later ISP adoption. We return to this issue when we address endogeneity in Column IV.

We now turn to the main results in the paper, which address all of the features mentioned above as well as the potential endogeneity of entry decisions. Again, the result of primary interest is the first row, which measures the effect of the expected number of competitors on technology s on the probability of adopting technology s. As stated in Section 6.2, we take two approaches to identifying the parameters of interest. The first, reflected in column III, is to assume that only ISPs that adopt the same technology affect profits on that technology. That is, row 3 (corresponding to ψ_3) is constrained to be 0. In the first row of column III, the coefficient is negative and significant. The effect is

	I	II	III	IV	V	VI		Test
# ISP's on same technology		-0.083	-0.047	-0.14	-0.011	-0.07	-0.02	0.24
		(0.008)	(0.015)	(0.07)	(0.005)	(0.03)	(0.02)	
# ISP's on both technologies		0.12	0.04	0.03	0.002	0.09	-0.02	0.21
		(0.01)	(0.03)	(0.03)	(0.008)	(0.06)	(0.02)	
# ISP's on other technology		-0.04		0.10				
		(0.01)		(0.07)				
Constant	-3.27	-5.40	-4.39	-4.35	-0.50	0.03	-7.78	0.01
	(2.05)	(0.43)	(0.68)	(2.11)	(0.61)	(3.00)	(0.88)	
ln of # of switches ISP covers	0.11	0.14	0.11	0.11	0.04	0.16	0.06	0.43
	(0.06)	(0.06)	(0.06)	(0.06)	(0.02)	(0.08)	(0.08)	
ln of # of switches squared	-0.01	-0.01	-0.01	-0.01	-0.004	-0.02	0.00	0.28
	(0.01)	(0.01)	(0.01)	(0.01)	(0.002)	(0.01)	(0.01)	
ISP has digital connection	0.66	0.63	0.65	0.66	0.19	0.72	0.58	0.25
	(0.05)	(0.05)	(0.05)	(0.06)	(0.02)	(0.08)	(0.09)	
Same technology had lead in July	-0.26			0.26				
	(0.06)			(0.20)				
ln(# of ISP's at switch)	0.03	0.13	0.11	0.09	0.02	0.12	0.06	0.57
	(0.04)	(0.04)	(0.05)	(0.05)	(0.01)	(0.07)	(0.08)	
# of backbone providers	-0.001	0.003	0.004	0.004	0.001	0.006	0.002	0.56
	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.005)	(0.004)	
ln(median HH Income)	0.17	0.34	0.26	0.27	0.05	-0.17	0.60	0.004
	(0.20)	(0.04)	(0.07)	(0.20)	(0.06)	(0.29)	(0.09)	
establishments per capita	4.00	15.30	6.75	7.73	1.89	-0.79	12.73	0.36
	(6.66)	(6.18)	(6.55)	(6.79)	(1.92)	(10.31)	(9.97)	
% pop college graduate	-3.46	-6.11	-3.74	-3.93	-1.03	-3.08	-3.48	0.90
	(1.69)	(1.20)	(1.31)	(1.73)	(0.50)	(2.54)	(2.04)	
% pop urban	0.05	0.03	0.02	0.02	0.00	0.35	-0.33	0.06
	(0.17)	(0.17)	(0.17)	(0.17)	(0.05)	(0.24)	(0.26)	
% pop in different cnty 5 yrs ago	1.18	1.95	1.27	1.27	0.30	1.96	0.34	0.14
	(0.48)	(0.47)	(0.49)	(0.50)	(0.15)	(0.72)	(0.78)	
Rho	-0.10	-0.22	-0.10	-0.10		-0.09		
	(0.04)	(0.05)	(0.04)	(0.04)		(0.04)		

Notes: Column I predicts adoption of X2 and Flex based only exogenous variables.
Column II includes choices of competitors but treats them as exogenous.
Column III treats competitor made in both July and October as symmetric endogenous variables.
Column IV treats decisions observed in July as exogenous to those in October.
Column V repeats column III but in a linear probability model estimated via TSLS. Instruments are predictions from Column 1.
Column VI repeats Column III but allows parameters to vary across Flex and X2 (sub-columns 1 and 2).
"Test" presents variable-by-variable Wald tests of equality for Column VI.
For each column, the number of observations is 2206. Standard errors are in parenthesis.

Table 5: Results

reasonable and important. Doubling the number of competitors adopting Flex reduces the probability of adopting Flex for a given ISP by 5.06 percentage points, a 16.4% change. These numbers are evaluated at means in the data, where the average number of competitors on Flex is 5.78 and the average number of competitors on X2 is 7.02. The coefficient is somewhat less negative than that in Column II, as would be expected if the main source of endogeneity is ISP-specific rather than location specific.

Our second identification strategy is to use adoption decisions in July as exogenous variation affecting ISP preferences over each technology. These results in appear in Column IV. Here again, the parameter in the first row is negative and significant at a 95% confidence level (although just barely – the t-statistic is 1.99). The parameter implies a larger marginal effect than the parameter in column III (although less precisely estimated). Doubling the number of competitors adopting Flex reduces the probability of adopting Flex for a given ISP by 11.01 percentage points, a 35.7% change.

The parameter on "both" is insignificant. We interpret this result as supporting the assumption that early adoption is exogenous. Because the parameter on "both" switches from large and significant to small and insignificant when correcting for endogeneity, it must be because the instruments do not capture whatever caused the endogenous variable to be large. If the parameter was large in Column II due to location specific factors, then the small coefficient when instrumenting with early adoption implies that early adoption does not reflect location specific factors. We take this result as evidence that ISP specific factors shaped early adoption much more than location specific factors, in which case early adoption is a valid instrument.

We also see from Column IV that the number of ISPs on the other technology is positive and not precisely estimated. This insignificant result may indicate that the number of ISPs on the other technology is not important and that it can reasonably be excluded, which would support the results in column III.

Accordingly, we consider the columns III and IV to be the main results in the paper. Both confirm our major hypothesis that ISPs preferred to differentiate from each other rather than coordinate on a single technology.

As a robustness check, Column V estimates an alternative functional form. We use a linear probability model in which endogeneity issues are addressed via two-stage least squares. The first stage are linear regressions predicting the expected number of competitors adopting the same technology, the other technology and both technologies. In addition to the firm's technological and average demographic characteristics, we use as instruments the expected number of adopters of each possibility as computed by Equation 4. The probabilities of each ISP adopting are constructed based on the results in Column I. That is, we use predictions of the number of adopters based only on exogenous variables as instruments in an otherwise linear procedure. The signs and significance tests are the same for the first two rows as in Column III and are similar for all other variables. We find the same results when re-estimating Column IV in a linear framework.

Column VI reestimates column III but allows parameters to differ across technologies. Again, the first row is negative although it is significant only for the Flex regression. We are primarily interested in these results in order to test whether the parameter restrictions are reasonable. Joint Wald or Likelihood ratio tests of equality are strongly rejected. However, the parameters appear broadly similar across specifications and a variable-by-variable Wald test shows that equality can be rejected at a 95% level of confidence for only two variables. Given there is strong institutional support for the symmetry of the technologies, we take this as reasonable evidence in favor of the restrictions imposed in the earlier regressions.[22] We find similar results when allowing parameters to differ

[22] We also tested a variety of specifications capturing whether this behavior varied across markets with different numbers of ISPs. However, after conditioning out differences in ISPs and local demographics we did not find large differences in behavior over markets of different

across technologies in a linear framework similar to that in Column V.

8 Sources of Adoption Delay

As it turned out, 56K modem sales to ISPs went very slowly relative to what the market could support. As we discussed, barely 50% of ISPs adopted 56K by October 1997. Furthermore, none of the large ISPs (AOL, AT&T, UUNET, MSN, GTE, Bell-South, EarthLink) adopted. Due to the large skew in market share (e.g. twenty ISPs served more than three quarters of the users), the vast majority of customers could not use 56K unless they switched from their existing large ISP to one of these smaller ISPs. Most consumers did not make this switch, even though most geographic regions had at least one or more ISP carrying 56K. Accordingly, sales to consumers were much less than what the potential market might have been.

This situation was soon remedied by the International Telecommunication Union (ITU) which announced the V.90 in February 1998. The ITU is an agency of the United Nations that brings government regulators and industry participants together to standardize telecommunication technology. The ITU is a voluntary standard setting organization. Unlike a regulatory organization, it requires a consensus to promulgate a standard. That is, it required the agreement of both the Rockwell and US Robotics consortia to move forward. In this case, a consensus was reached on a new standard that had the same functionality but was incompatible with both X2 and Flex. Despite the fact that the ITU had no enforcement power in this case and served only to create a focal point, the V.90 was remarkably successful and led to fast and widespread adoption by both consumers and ISPs.[23]

size, as Figure 2 suggests.

[23] While market participants report sales increasing with the introduction of the V.90, data is difficult to come by. One example comes from the Graphics, Visualization and Usability laboratory at the Georgia Institute of Technology which has conducted a WWW users survey

What role did ISP differentiation have on the initial coordination failure via the market and the eventual successful coordination by committee? Why did we observe a costly standards war when a standard setting organization existed to prevent its occurrence? Greenstein and Rysman (2004) provide a case study of the relationship between ITU negotiations and market outcomes in the case of 56K modems. Here, we summarize some of those results and relate them to ISP differentiation. Our discussion in this section is necessarily more speculative than the empirical work that has preceded it.

Greenstein and Rysman (2004) show that most important industry participants believed that the ITU would eventually produce a standard and that the market would coordinate on that standard. However, the timing and features of the standard was very difficult to predict. Many participants believed that the relative success of the two technologies in the market place might shape the eventual outcome at the ITU.

Market outcomes seem to have primarily affected the timing of the new standard. Meetings began at the ITU by December 1996, before product introduction. The ITU promised a standard within two years. However, two years would be quick relative to previous ITU decisions. Farrell (1996) reports that similar organizations deliver standards in 5 years on average. In fact, the ITU announced the V.90 standard in only 18 months. At the time, this was regarded as the shortest period of time the ITU had ever required to reach a decision. (Press Release ITU/98-4). Certainly, the timing of the ITU agreement is striking in that it came soon after the November announcements about adoption decisions for the largest ISPs to split across technologies.

semiannually since January 1994. The surveys cover a broad range of topics but one portion of the survey inquires about speed of connection to the Internet. The GVU survey of on-line users from April 1997 and October 1997 showed that the percentage of dial-up users with 56K modems was, respectively, 362 out of 8897 and 377 out of 3822, which is less than ten percent. In April and October of 1998 the percentages are much higher, respectively, 1242 out of 4654 and 760 out of 1534, almost a quarter and almost one half.

How might differentiation by ISPs have affected this process? ISP differentiation and low ISP adoption rates by themselves could have raised the benefits to quick agreement at the ITU. Furthermore, if ISP differentiation was a source of consumer delay, then ISPs may have played an even larger role in determining the progress of the market. Rysman (2003) provides a formal model of how ISP differentiation might lead to adoption delay by consumers. The model has two equilibria. In one, ISPs and consumers believe that the technology that achieves a small early lead (even if from a random event) will become the market leader and jointly adopt that standard. In the second equilibrium, ISPs believe that consumers will delay their adoption until one technology achieves a large lead. ISPs that adopt a less popular standard have less of a chance of serving a large market in the future but earn higher margins on early adopters because they face less competition. If the effects of competition are large, ISPs prefer the short-run benefits of less competition to the long-run benefits of supplying a large market. In this case, consumers know a small early lead does not indicate a technology will become a standard and rationally delay their adoption decision until one technology achieves a sizeable lead. Eliminating one of the technologies leads to immediate widespread adoption. The goal of the model is to associate consumer delay with ISP differentiation and show how adoption is accelerated when only one standard exists.

An opposing explanation of consumer delay is that many consumers were waiting for the ITU standard and would have delayed their adoption regardless of ISP decisions. However, the technology sponsors seem to have thought they could affect consumer purchasing before the ITU decision, and they also seem to have calculated that competing fiercely in the early deployment was optimal in case the standard emerged slowly from the ITU. The sponsors of the two technologies engaged in a "standards war," promoting their preferred standard with discounts as well as marketing promotions aimed at growing the size of the consortia associated with their technology. Much of this activity from sponsors was aimed at persuading non-adopters that a sponsor's technology would be the widely adopted standard. Relatedly, some of the trade press questioned whether the ITU would produce a useful standard in any reasonable time.

The practice of bundling modems with computers does not provide an explanation of consumer behavior. The external modem market was much larger than the internal modem market at this time. Furthermore, computer manufacturers did not align themselves with one technology or the other. It seems that computer assemblers were beset by the same confusion as consumers, which was presumably due in some part to ISP differentiation. While many forces were at work overall, the 56K modem story is consistent with one in which adoption delay and ISP differentiation reinforced each other.

Our broad point is that the events during early deployment highlight the nuanced intersection between market behavior and the deliberations of standard setting organizations, such as the ITU. If the deployment of 56K went differently (if for instance, ISPs preferred coordination to differentiation), so too would have the negotiations at the ITU. In that sense these market events altered the willingness of firms to participate in ITU deliberations, and it affected their stance in the consensus process shaping the final specification. It also altered the willingness of firms to delay on settlement or settle quickly. Hence, this case illustrates how success (or lack of it) in the market place can take on crucial importance in determining the timing of resolution, as well as the actual specification chosen under negotiated settlement.

9 Conclusion

This paper studies the importance of competition in technology adoption. We exploit a unique data set on the standardization process for 56K modems in numerous geographically independent markets. We show that Internet Service Providers split evenly across the two available standards in local markets, confirming the importance of competition. We show that ISPs split much more evenly than independent random choice would predict. We confirm this result in a bivariate probit framework that controls for ISP characteristics, demographics and decision-making by ISPs in multiple markets. Finally, we verify the result in a model based on an entry game of imperfect information that controls for the endogeneity of entry between rival ISPs. The fact that competitive forces are so strong is particularly surprising given the presence of an indirect network effect between ISP and consumer adoption of a 56K standard.

In a more speculative section, we discuss the role of ISP differentiation in the coordination failure in the market with two standards and the eventual successful coordination by the market on a new standard introduced by a voluntary standard setting organization. While standard models of network effects might predict rapid "tipping" in a competition between homogenous standards, we point out that service provider competition provides a force in the opposite direction. To the extent that ISP differentiation was associated with consumer delay, it may have led the technology sponsors towards a quick agreement and an end to the costly standards war.

10 Appendix

10.1 Multinomial Test of Agglomeration and Dispersion

This subsection presents functional forms for the Multinomial Test of Agglomeration and Dispersion (MTAD). A fuller description may be found in Rysman and Greenstein (2004). Suppose we observe M markets each populated by n_m agents $m = 1, ..., M$, where n_m is bounded by $\underline{n} > 0$ and $\overline{n} < \infty$. The variable n_m is distributed according to the discrete distribution $f(n_m)$. The agents choose between C options, available in each market. The unconditional probability of observing option c is p_c, $c = 1, ..., C$. The observed number of agents choosing option c in market m is x_m^c. Let $\mathbf{x}_m$ be the vector of elements $x_m^1, ..., x_m^C$ and $\mathbf{p}$ be the vector of probabilities $p_1, ..., p_C$. Let $\mathbf{X}$ be the $M \times C$ matrix of choices for all of the markets and $\mathbf{n}$ be the $M \times 1$ vector of the number of agents in each market. If the agents make choices independently, the average log likelihood of observing the outcome $x_m^1, ..., x_m^C$ in for M markets is:

$$l(\mathbf{X}, \mathbf{n}, \mathbf{p}) = \frac{1}{M}\sum_{m=1}^{M} \ln\left(\binom{n_m}{x_m^1, ..., x_m^C}\right) + x_m^1 \ln(p_1) + ... + x_m^C \ln(p_C)$$

Consider the likelihood value if the data were actually generated by independent random choice. Let the random variable $l(f, \mathbf{p})$ be distributed according to the distribution $l(\mathbf{X}, \mathbf{n}, \mathbf{p})$ if $\mathbf{X}$ was actually drawn from a multinomial distribution and n_m was drawn from f. Then we have that:

$$E[l(f, \mathbf{p})] = \sum_{n_m=\underline{n}}^{\overline{n}} \sum_{\mathbf{z}\in\Xi(n_m)} \left(\ln\left(\binom{n_m}{z^1, ..., z^C}\right) + z^1 \ln(p_1) + ... + z^C \ln(p_C)\right) L(\mathbf{z}, n_m, \mathbf{p}) f(n_m)$$

where $\Xi(n_m)$ is the set of all possible choice configurations of n_m agents. Our test statistic, which can be show to converge to zero with normal variance under the null is:

$$t(\mathbf{X}, \mathbf{n}, \mathbf{p}) = l(\mathbf{X}, \mathbf{n}, \mathbf{p}) - E[l(f, \mathbf{p})]$$

10.2 Computing the fixed point algorithm

We can write $E[\Pi_j^s]$ as:

$$E[\Pi_j^s] = x_j^f \beta_2 + \frac{1}{I}\sum_{i\in\vartheta_j} x_i^l \beta_1 + \psi_1 \frac{1}{I}\sum_{i\in\vartheta_j} E[n_i^s|\mathbf{X}, \theta] + \psi_1 + \psi_2 \frac{1}{I}\sum_{i\in\vartheta_j} E[n_i^{AB}|\mathbf{X}, \theta]$$

Let $\Lambda(j)$ be the set of firms that overlap with j and let λ_{jk} be the number of switches at which j overlaps with k. Then we have that:

$$\sum_{i\in\vartheta_j} E[n_i^s|\mathbf{X}, \theta] = \sum_{i\in\vartheta_j}\sum_{k\in\Upsilon(i)} P_{ks} = \sum_{k\in\Lambda(j)} \lambda_{jk} P_{ks}$$

Therefore, the average expected number of entrants faced by firm j depends only on the number of locations at which j faces entrants and their entry probabilities, not on where j faces them. That is, facing 3 markets with one potential entrant in each has the same competitive implications as facing 2 markets with no potential entrants and 1 market with 3, assuming entry probabilities are the same. This feature allows us to solve the fixed point algorithm in an economical manner. Let $\mathbf{\Lambda}$ be an $n \times n$ matrix with λ_{jk} in position j, k with $\lambda_{jk} = 0$ if $j = k$. Let ψ be the 4×1 matrix with elements 1,2,3 and 4 measuring the competitive impact of firms adopting none, the same standard, the other standard and both. In our implementation, $\psi = [0, \psi_1, 0, \psi_2]$ for Flex and $\psi = [0, 0, \psi_1, \psi_2]$ for X2. Then:

$$E[\Pi_j^s] = x_j^f \beta_2 + \frac{1}{I}\sum_{i\in\vartheta_j} x_i^l \beta_1 + \mathbf{\Lambda P}\psi + \psi_1 \qquad (5)$$

Solving the fixed point algorithm consists of picking a starting matrix $\mathbf{P}^0$, plugging into Equation 5 and plugging the result into Equation 2 to obtain a new matrix of adoption probabilities $\mathbf{P}^1$. We can iterate this way until convergence. The term $\sum_{i\in\vartheta_j} x_i^l/I$ can be computed once at the beginning of the estimation algorithm, further reducing the computational burden.

Note how restricting the competitive impact to be linear reduces the computational burden: we need only know the number of locations at which any firm overlaps with any other firm. If we were to add square terms, we would need to know the number of locations at which a firm overlaps with any pair of competitors and cubic terms require us to know the how many locations a firm overlaps with any triplet. Putting in a more complicated term, such as the ratio of adopters of one standard to the other, would likely requires us to loop through every firm-location combination in the data (more than 200,000 items) and would be infeasible to our knowledge. It is straightforward to add these terms to the analysis that assumes exogeneity and we found that it did not change the analysis. For instance, adding squares of the number of competitors adopting one of the choices does not change any of the results in a substantial way.

References

[1] Ackerberg, D and G. Gowrisankaran (2004) "Quantifying Equilibrium Network Externalities in the ACH Banking Industry," *mimeo*, UCLA

[2] Augereau, A, and S. Greenstein (2001)," The Need for Speed in Emerging Communications Markets: Upgrades to Advanced Technology at Internet Service Providers." *International Journal of Industrial Organization*, v 19, pp 1085-1102.

[3] Besen S.M. and L. Johnson (1986), *Compatibility Standards, Competition, and Innovation in Broadcasting Industry*, Rand Report, R-3453/1-Nsf.

[4] Berry, S.T. (1992) "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, 60(4), 889-917.

[5] Church, J. and N. Gandal (1992) "Network Effects, Software Provision and Standardization," *Journal of Industrial Economics*, 40(1) 85-103.

[6] Chou, C. and O. Shy (1990) "Network Effects without Network Externalities" *International Journal of Industrial Organization*, 8 259-270.

[7] Dranove, D and N. Gandal (2003) "The DVD vs. DIVX Standard War: Empirical Evidence of Network Effects and Preannouncement Effects,"*Journal of Economics and Management Science.*, 12 (3), pp. 363-386.

[8] Downes, T. and S. Greenstein (1999), "Universal Access and Local Commercial Internet Markets "*Research Policy*, 31, 1035-1052.

[9] Ellison, G. and E. Glaeser (1997) "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach," *Journal of Political Economy*, 105(5) 889-927.

[10] Farrell, J. (1996) "Choosing Rules for Formal Standardization," mimeo, University of California, Berkeley.

[11] Gandal, N., M. Kende and R. Rob (2000), "The Dynamics of Technological Adoption in Hardware/Software Systems: The Case of Compact Disc Players", *Rand Journal of Economics*, 31, 43-61.

[12] Genesove D, (1999) "The Adoption of Offset Presses in the Daily Newspaper Industry in the United States," NBER Working Paper 7076.

[13] Gowrisankaran, G. and J. Stavins (2004) "Network Externalities and Technology Adoption: Lessons from Electronic Payments," *RAND Journal of Economics*, 35, 260-276.

[14] Greene, T (1997) "More Bumps in the Road for 56K modems," *Network World*, September 22, 1997, p12.

[15] Greenstein, S. (2000) "Building and Delivering the Virtual World: The Commercial Internet Access Market," *Journal of Industrial Economics,* 48(4), 391-411.

[16] Greenstein, S. and M. Rysman (forthcoming) "Lessons about Standard Setting from Modems," *Standards and Public Policy* Editors, S. Greenstein and V. Stango, Cambridge University Press.

[17] Haile, P.A., A. Hortascu and G. Kosenock (2003) "On the Empirical Content of Quantal Response Equilibrium," mimeo, Yale University.

[18] Klepper, S. (2002) "Firm Survival and the Evolution of Oligopoly," *RAND Journal of Economics*, 33(1) 37-61.

[19] Manski, C. (1993) "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 60, 531-542.

[20] Mazzeo, M (2002) "Product Choice and Oligopoly Market Structure," *RAND Journal of Economics,* 33(2) pp 1-22.

[21] McKelvey, R.D. and T.R. Palfrey (1995) "Quantal Response Equilibrium for Normal Form Games," *Games and Economic Behavior,* 10(1), 6-38.

[22] Mulligan, J.G. and E. Llinares (2003) "Market Segmentation and the Diffusion of Quality-Enhancing Innovations: The Case of Downhill Skiing," *Review of Economics and Statisitcs*, 85(3), 493-502.

[23] National Telecommunications Information Administration (2002),. "A Nation Online: How Americans Are Expanding Their Use Of The Internet.' http://www.ntia.doc.gov/reports.html.

[24] Ohashi, H (2003) "The Role of Network Effects in the US VCR Market, 1978-1986," *Journal of Economics and Management Strategy* 12(4), 447-494.

[25] Park, S (2000) "Quantitative Analysis of Network Externalities in Competing Technologies: The VCR Case," *mimeo*, SUNY-Stony Brook.

[26] Postrel, S (1990) "Competing Networks and Proprietary Standards: The Case of Quadraphonic Sound," *Journal of Industrial Economics,* 39(2), 169-185.

[27] Rickard, J (1998) "The 56K Modem Battle," *Boardwatch*, March 1998.

[28] Rickard, J (1997a) "56K Modems: The Battle Continues," *Boardwatch*, March 1997.

[29] Rickard, J (1997b) "U.S. Robotics Launches the New Battle - 56Kbps Modems" *Boardwatch,* January 1997.

[30] Rohlfs, J (2001) *Bandwagon Effects in High Technology Industries*, MIT Press, Cambridge.

[31] Rysman, M. (2003), "A Standards War Under Indirect Network Effects," *mimeo*, Boston University.

[32] Rysman, M. (2004), "Competition Between Networks: A Study of the Market for Yellow Pages," *Review of Economic Studies*, 71(2) 483-512.

[33] Rysman, M. and S. Greenstein (2004), "Testing for Agglomeration and Dispersion," *Economics Letters,* forthcoming.

[34] Saloner, G. and S. Shepard (1995) "Adoption of Technologies with Network Effects: An Empirical Examination of the Adoption of Automated Teller Machines" *RAND Journal of Economics*, 26, 479-501.

[35] Shapiro, C., and H. Varian (1999), *Information Rules: A Strategic Guide to the Network Economy*, Harvard Business School Press, Cambridge.

[36] Shy, O. (2001) *The Economics of Network Industries*, Cambridge University Press, Cambridge.

[37] Seim, K. (2004) "An Empirical Model of Firm Entry with Endogenous Product-Type Choices," mimeo, Stanford University GSB.