Low-Income Demand for Local Telephone Service: Effects of Lifeline and Linkup*

Daniel A. Ackerberg
University of California at Los Angeles

Michael H. Riordan
Columbia University

Gregory L. Rosston
Stanford Institute for Economic Policy Research

Bradley S. Wimmer
University of Nevada, Las Vegas

September, 2009
REVISION

Abstract

This policy study evaluates the effects of the "Lifeline" and "Linkup" subsidy programs on telephone penetration rates of low-income households, and provides a framework for evaluating similar policies for Internet access. These two programs, respectively, subsidize the monthly subscription and initial connection prices for low-income households. Our demand specifications use location-specific subsidized prices and a discrete choice model aggregated across demographic groups. GMM estimators correct for endogeneity of these prices. Our policy simulation suggests that penetration rates would be 4.1 percentage points lower without the policies, that Linkup is more cost-effective than Lifeline, and that automatic enrollment policies are important.

*The authors thank Mitali Das, Gautam Gowrisankaran, Jinyong Hahn, Katherine Ho, and seminar participants at Columbia, Harvard, Melbourne, and MIT for helpful comments, Jisun Baek , Anna Bordon, David DeRemer, Carlos Perez, Ting Wu, Dan Steinert, Steve Loughlin, Jason May, Adrian Cacuci and Caroline Sandifer for excellent research assistance, the staffs of public utility commissions for guidance about the local rates and low-income programs in their respective states, and Jim Lande and Tracy Waldon for information on data available from the FCC. An early version of this paper was presented at the 2003 Telecommunications Policy Research Conference.

Abstract

This policy study evaluates the effects of the "Lifeline" and "Linkup" subsidy programs on telephone penetration rates of low-income households, and provides a framework for evaluating similar policies for Internet access. These two programs, respectively, subsidize the monthly subscription and initial connection prices for low-income households. Our demand specifications use location-specific subsidized prices and a discrete choice model aggregated across demographic groups. GMM estimators correct for endogeneity of these prices. Our policy simulation suggests that penetration rates would be 4.1 percentage points lower without the policies, that Linkup is more cost-effective than Lifeline, and that automatic enrollment policies are important.

Introduction

Measuring the effectiveness of government programs is important not only for evaluating existing policies, but also for contemplating the continuation or expansion of the policies going forward. A case in point is telecommunications universal service policy. Universal service for telephony has at least nominally been a public policy concern for quite a while. Universal service policies for ordinary telephone service were expanded significantly in the wake of the 1996 Telecommunications Act, subsequently were expanded to encompass wireless service, and currently are under debate for Internet access service. Universal service concerns usually are directed at two different, but somewhat overlapping, groups: rural and low-income households. Our focus is to develop a model for the demand of low-income households and the economic factors affecting their decisions to subscribe to telephone service. Our model evaluates the effectiveness of the post-1996 Lifeline and Linkup subsidy programs at increasing the telephone penetration of low-income households. More generally, our study develops an appropriate methodology and collects appropriate data for evaluating the effectiveness of low-income subsidy programs. Such a framework, and an understanding of its data requirements, is important for evaluating current telephone subsidy programs, and potentially for the gathering debate on Internet access subsidies. In fact, legislation has just been introduced to expand the Lifeline subsidy program to cover broadband adoption (Office of Congresswoman Doris Matsui, 2009).

Overall telephone penetration in the U.S. is high – almost 95% according to the Federal Communications Commission (FCC) "Penetration Report." (Belinfante, 2007) This same report shows that penetration rates are significantly lower for low-income households. Less than 92.3% of households with income less than $20,000 had a working telephone in their households.

Disaggregated data are not available in the most recent report, but the 2003 Penetration Report (Belinfante, 2003) shows that low-income penetration rates differ substantially across states, ranging from 96.5% in Maine (compared to 98.2% of all households) to under 80% in Mississippi (compared to 90.9% of all households).

The Lifeline program, started in 1985, provides a subsidy that reduces monthly charges for eligible low-income subscribers. The Linkup program reduces the initial connection fee that low-income households pay to establish telephone service. States may augment these federal programs.

In 1996, the FCC dramatically increased the size of its Lifeline subsidy. Prior to the Telecommunications Act of 1996 ("Act"), the federal Lifeline program waived the federal subscriber line charge (SLC), which was equal to $3.50 per month in most jurisdictions, as long as states matched this with a state-funded subsidy to low-income households. The post-Act Lifeline program provided all low-income customers in all states a baseline support equal to the federal SLC plus $1.75, for a total of $5.25 in all jurisdictions in 2000 with the exception of the District of Columbia where the federal SLC was less than $3.50.[1] Lifeline customers received additional federal support equal to one half of any support provided by an intrastate program, up to $7.00 in total federal support. In states that took full advantage of the matching federal program Lifeline customers receive a subsidy of $10.50.[2] The federal Linkup program reduces low-income subscribers' initial connection charge by 50 percent of the customary charge, or $30, whichever is less.[3] The FCC's implementation of the 1996 Act did not change the federal Linkup subsidy.

Several studies have examined the effect that Lifeline and Linkup programs have on penetration rates.[4] The majority of studies used state-level data that include measures of the size of Lifeline and Linkup programs as explanatory variables in regressions that estimate the overall penetration rate in a state. For example, Garbacz and Thompson (2002, 2003) used state-level

---

[1] The size of the baseline support has risen in recent years because the cap on the federal SLC for residential customers has increased.
[2] States have the ability to provide additional unmatched subsidy as well.
[3] Both Lifeline and Linkup are funded by taxes on telecommunications services. To the extent that low-income households are heavy users of the services taxed (e.g. long distance), the overall price reduction is less. We recognize that marginal subscribers are not likely to be heavy users of the taxed services, so low-income telecommunications users presumably experience a price decrease. See Hausman, Tardiff and Belinfante (1993).
[4] See Riordan (2002) for a more complete background on the economics of universal service.

data from the 1970, 1980, 1990 and 2000 decennial censuses to estimate penetration rates. Erikson, Kaserman and Mayo (1998) used state-level data from the current population survey ("CPS") to conduct their study. Both studies found that Lifeline and Linkup programs have a statistically significant but small impact on overall penetration rates. Garbacz and Thompson (2003) found that the demand for local service is highly inelastic (-0.006 to –0.011 in 2000) and that Lifeline and Linkup programs had little effect on penetration, estimating that a 10 percent increase in Lifeline and Linkup expenditures would have added only about 20,000 households to the network in 2000. Erikson, Kaserman and Mayo (1998) found that targeted low-income subsidies affect state-level penetration rates positively, while untargeted subsidies do not have a statistically significant impact on penetration rates.

Studies that rely on statewide data use statewide-average residential prices as an independent variable. Because residential service prices can vary substantially within states, the use of statewide data masks substantial information. For example, in California in 2000, monthly rates for 100 calls a month for Lifeline customers vary from $5.01 to $6.90 and for non-Lifeline customers vary from $11.62 to $15.51.

Crandall and Waverman (2000) used location-specific price data obtained from incumbent local exchange carriers (ILECs) and data from the 1990 census for 1,897 towns or places. They used the price of local service and the Lifeline rates as alternative explanatory variables, as well as a dummy variable measuring whether the state has a Lifeline or Linkup plan interacted with the poverty rate, to try to measure whether poor communities in states with Lifeline and Linkup programs have higher penetration rates than poor communities in other states. They found no significant effect of Lifeline programs, which is consistent with their finding of little price elasticity of demand for telephone service overall. Crandall and Waverman found that a higher charge for connecting a new subscriber reduces penetration rates, estimating an elasticity of penetration with respect to the connection charge ranging from -0.025 to –0.030. Surprisingly, a Linkup program was associated with lower penetration rates. Crandall and Waverman explain that the counterintuitive Linkup effect may be due to their use of a dummy variable for Linkup and the fact that only two states did not have a Linkup program in 1990. They also suggest that the Linkup result may be due to a reverse-causation problem. Specifically, states that have high penetration rates may choose not to participate in federal low-income programs. Earlier studies had estimated higher price elasticities of demand for low-income households than the average income household, thus providing some empirical justification for the first versions of the Lifeline and Linkup programs that were implemented in the 1980's (Perl, 1984; Taylor and Kridel, 1990; Cain and McDonald, 1991).

Our study differs from previous work in at least three important ways. First, using various data sources, we have constructed a dataset that is more extensive than other datasets used to study low-income telephone penetration. Unlike Garbacz and Thompson (2002, 2003), we have prices at a much more disaggregate level. Unlike Crandall and Waverman (2000), we consider penetration specifically for poor households (rather than overall penetration), and we have specific Linkup prices rather than a Linkup dummy. One virtue of restricting attention to poor households is that implicitly we allow price sensitivity for low-income populations to differ from the rest of the population. Another is that we can directly exploit price variation resulting from new Lifeline subsidies introduced in wake of the Act. Second, our preferred specification controls for the possible endogeneity of Lifeline prices. Lifeline price endogeneity is a concern because states responded to post-1996 changes in federal Lifeline policy differently. Ignoring this endogeneity potentially biases downward the estimated elasticity of demand with respect to Lifeline prices. Finally, we also use the size of the local calling area as an explanatory variable.[5] The inclusion of this value-of-service variable in the demand specification by itself limits price endogeneity because states typically set higher prices in places with larger local calling areas.

Our empirical analysis uses connection and monthly subscription prices for households eligible for Lifeline and Linkup programs, and the characteristics of relevant service plans. Data on prices and service characteristics obtained from Bell Operating Company (BOC) tariffs, and Census data on telephone penetration and demographics, are matched to more than 7,000 wire centers. A wire center is a geographic area that includes all customers connected to a particular local switch. Since the Census data only reports the aggregate low-income penetration rate at each location, our empirical specifications are based on an underlying discrete choice model of household demand for telephone service that is aggregated across demographic groups using information on the demographic composition of each location. The rich dataset and our exclusive focus on poor populations allow us to estimate elasticities with respect to Lifeline and Linkup prices. These elasticity estimates are employed to evaluate the effectiveness of Lifeline and Linkup programs at increasing low-income telephone penetration.

[5] Perl (1984) and Taylor and Kridel (1991) use the size of the local calling area as an explanatory variable.

Our main conclusion is that Lifeline and Linkup subsidies increased the telephone penetration of poor households in our sample by 4.1 percentage points, with a 95% confidence interval between 1.6 and 6.8 percentage points. We also find that Linkup is more cost effective than Lifeline because it is targeted at low-income households that do not have telephone service, and that the automatic enrollment programs are effective at increasing telephone penetration.

Theory and Empirical Specification of Household Telephone Demand

Telephone service enables a household to place and receive calls. The value of telephone service to a representative household is assumed to be multiplicatively separable in the characteristics of the household and the characteristics of the service. Specifically, we assume a household is willing to pay $\phi(e^{t+V})$ for telephone service, where t describes the household, V describes the service, and $\phi(\bullet)$ is a strictly increasing function. If the price of telephone service is R, then the household elects service if $e^{t+V} \geq \phi^{-1}(R)$, or equivalently, if $t \geq \ln\phi^{-1}(R) - V \equiv \psi(R) - V$, where $\psi(\bullet)$ is a strictly increasing function.

Consider a population of households described by a cumulative distribution function $F(t)$. The share of households who demand the service (penetration rate) at price R is

$$\overline{S} = 1 - F(\psi(R) - V)$$

Next, partition the population into M demographic groups, indexed $g = 1,...M$. Let X_g denote the population share of group g, and $F_g(t)$ the distribution of t for group g. Then telephone penetration of group g is

$$\overline{S}_g = 1 - F_g(\psi(R) - V)$$

and the penetration rate of the whole population is

$$\overline{S} = \sum_{g=1}^{M} X_g \overline{S}_g$$

Now suppose there is a finite population of households of size N with group shares $(X_1,..X_M)$. Interpret $\overline{S}_g$ as the probability that a randomly chosen member of group g adopts, and $\overline{S}$ as the probability that a randomly chosen member of the whole population adopts. Thus the realized number of households in group g with telephone service is a draw from a binomial distribution $\mathrm{B}(\overline{S}_g, X_g N)$, and the total number of households with service is a draw from the convolution of the distribution functions for the *M* groups.

Finally, assume the distribution of household types for any group is exponential, with $F_g(t) = 1 - e^{-\lambda(t-\mu_g)}$ for $t \geq \mu_g$, where μ_g is a group-specific location parameter and λ is a common scale parameter. Assume that no group is certain to have 100% penetration, i.e. $\overline{S}_g < 1$ for all g (this is the case if $\mu_g \leq \psi(R) - V$ for all g). It follows that $\overline{S}_g = e^{-\lambda(\psi(R)-V))+\lambda\mu_g}$ and $\overline{S} = e^{-\lambda(\psi(R)-V))} \sum_{g=1}^{M} X_g e^{\lambda\mu_g}$. Alternatively, the expected penetration of the entire population is explained by the logarithmic equation

$$\ln \overline{S} = -\lambda\psi(R) + \lambda V + \ln\left(\sum_{g=1}^{M} e^{\lambda\mu_g} X_g\right) \tag{1}$$

This special case provides the basis of our first approach to estimation.

We can allow $\overline{S}_g = 1$ for some groups by proceeding more generally and letting

$$I_g = \begin{cases} 1 & \text{if } \mu_g \leq \psi(R) - V \\ 0 & \text{otherwise} \end{cases}$$

If $I_g = 0$ then telephone service is sufficiently valuable that all members of group *g* adopt it. Then $\overline{S}_g = \min\left\{e^{-\lambda(\psi(R)-V))+\lambda\mu_g}, 1\right\}$ and

$$\ln\left(\overline{S} - \sum_{g=1}^{M} (1 - I_g) X_g\right) = -\lambda\psi(R) + \lambda V + \ln\left(\sum_{g=1}^{M} e^{\lambda\mu_g} I_g X_g\right) \tag{2}$$

This more general model of the expected penetration rate underlies our second approach to estimation.

The second approach has advantages and disadvantages compared to the first approach. On one hand, the second approach allows particular groups in particular locations to adopt with probability one. On the other hand, as will be detailed below, the second approach results in a model that is not invertible in the econometric residual. This requires additional assumptions. Below we explain this distinction further.

A basic unit of observation is a population of consumers at location *l*. A vector of group shares $\left(X_{1l}, \ldots X_{Ml}\right)$ describes each population. Both the price (R_l) and other service characteristics (V_l) vary across locations. Our empirical model allows for variation in population and service characteristics across locations. We include in V_l any location specific variables that shift the distribution of tastes, and treat $(\mu_1,...\mu_M)$ as a fixed parameter vector. A simplifying assumption is that all population groups have the same scale parameter λ.

Our empirical analysis considers two alternative specifications of the function $\psi(\cdot)$: a linear model with $\psi(R) = R$, and a logarithmic model with $\psi(R) = \ln R$. In the linear model, consumer willingness to pay is $t+V$, and the price elasticity of demand is λR. In the logarithmic model, willingness to pay is $\ln(t+V)$ and the price elasticity is λ. We focus on the linear model, and treat the logarithmic model as a robustness check.

Our empirical model must deal with the fact that telephone service typically requires a monthly subscription price (P_l), and a one-time connection charge (C_l). If the household monthly "discount rate" is α, then $R_l = P_l + \alpha C_l$. The discount rate converts the one-time installation charge into a monthly household expense. We assume that this discount rate is constant.

It also is important to control for differences in the nature of service or distribution of tastes at different locations, such as the number of people within the local calling area. Let

$$\lambda V_l = \Gamma(Y_l) + \xi_l$$

where Y_l is a vector of observed characteristics at location *l* and unobserved characteristics are summarized by ξ_l. We assume that $\Gamma(\cdot)$ is linear in appropriately defined variables, and discuss possible distributional assumptions on ξ_l later. The unobservable ξ_l can also be interpreted to include a location-specific demand shock for telephone service.

Data

We composed our dataset using various sources: the 2000 decennial Census (United States Department of Commerce, 2000), BOC state telephone tariffs, the FCC, Telcordia (2000) (the Local Exchange Routing Guide, "LERG"), and Claritas (2003). Our unit of observation is the wire center. For each wire center, our data includes telephone penetration rates, demographics, and prices of basic local telephone service including connection charges, Lifeline and Linkup discounts, and other tariff information. In addition, we have variables that proxy the cost of providing local service and several other variables relevant for state regulation. These variables are used as instruments to control for possible price endogeneity. The basic data set includes 7,938 wire centers located in 43 states and the District of Columbia in the original BOC regions,[6] representing over 80 million residential access lines. FCC (2000a, 2000b) data indicate that the Lifeline program subsidized approximately 5 percent of the residential lines in our data in 2000.[7]

We collected data on local prices and other data from state tariffs at the level of what the LERG calls "localities," and used the LERG to match prices to wire centers. In metropolitan areas several wire centers serve a single locality, while in rural areas, a single switch may serve multiple localities. We dropped wire centers serving multiple localities with different prices from the sample, reducing the sample to 7,117 wire centers.[8] We matched wire centers to the census data using Claritas (2003).[9] The dependent variable is a wire center's penetration rate for low-income households (*Penetration*); this variable equals the number of low-income households with telephone service divided by the total number of low-income households. For our purposes, a low-income household is one below the poverty line.

The independent variables of primary interest are the monthly charge for local service and the connection charge for initiating service. Because low-income households are the focus of this study, we use Lifeline and Linkup rates for estimation. There is a potential problem in

[6] Excluded states are Alaska, Hawaii and Connecticut, which were not served by BOCs, Delaware, which is not included in the FCC (2000a) cost model, and Montana, Wyoming and Vermont, which set different prices for households served by each switch depending on the distance from the switch so that it was impossible to accurately determine the prices faced by low income households. Southern New England Telephone, which serves Connecticut, was purchased by SBC following passage of the Telecom Act of 1996.

[7] The FCC (2003) estimates that over 17 million households were eligible for the Lifeline program in 2000. Using FCC data on the percentage of households in each state that are eligible for the Lifeline program, we estimate that our data set includes approximately 8.5 million eligible households, or about half of all eligible households. We estimate that nearly 70 percent of the eligible households reside in areas included in this study (Alaska, Connecticut, Delaware, Hawaii, Montana, Vermont, and Wyoming are excluded from the study. In the remaining states we include only wire centers served by the BOCs where we can uniquely determine the price.) This study also does not include eligible households whose incomes exceed the poverty level (e.g., California households with incomes up to 150% of the poverty line are eligible for Lifeline and Linkup subsidies). We estimate that there are over 3 million eligible households that are above the poverty level residing in wire centers included in this study.

[8] In 2000, the FCC reported that only 47 percent of Indian tribal households on reservations and other tribal lands have a telephone. It provided additional Lifeline and Linkup monies to tribal areas and changed various other rules to promote subscribership in these areas. FCC (2000c). As a result, we drop one wire center in New Mexico because it contained a large proportion of Native Americans.

[9] Claritas (2003) cross references census block groups (CBGs) and wire centers. Census blocks include approximately one square block in metropolitan areas, although they are typically larger in rural areas. A typical CBG includes four or five CBs. In its publicly available data, the Census provides information on the availability of telephone service in low-income households for census tracts, which typically include four or five CBGs. In cases where a census tract crossed the wire-center boundaries, we allocated census data to a wire center based on the number of poor households in a CBG. We exclude census tracts where the census tract is served by multiple wire centers and the Lifeline rates are not the same.

doing this because many low-income households that are eligible for the subsidized rates actually pay the normal rates.[10] Presumably this is because of informational issues or the implicit transaction (and non-pecuniary) costs a household incurs to apply for the subsidized rates. However, as long as we assume that these non-adopters are not marginal consumers, our model is internally consistent.[11]

In the majority of states, Lifeline customers can choose from a variety of local-service offerings. Customers may subscribe to a usage-based plan, where they pay for each call or minute of local use in addition to a monthly charge.[12] Customers subscribing to a flat-rate plan pay a monthly charge and are allowed to make an unlimited number of local calls. The majority of the states in the sample offer both flat-rate and usage-based plans. Only Wisconsin and portions of New York (NYC) and Illinois require that consumers subscribe to a usage-based plan, while Kansas, Kentucky, North Carolina and Maine do not offer usage-based options.[13] The empirical analysis uses the variable *Lifeline50*, which is the minimum monthly expenditure of Lifeline customers making 50 local calls. As a robustness check, we also consider *Lifeline100*, which is the minimum monthly expense for making 100 calls.[14] The other price variable of primary interest is *Linkup*, which is equal to the connection charge paid by customers eligible for the Linkup subsidy.

The FCC reports that penetration rates for Black and Native American populations generally are lower than average while those for Asians are higher (Belinfante, 2007). To control for possible ethnic differences in the demand for telephone service we consider the

---

[10] The FCC (2003) estimates that in 2000 only 37.5 percent of the Lifeline-eligible households participated in the program. Burton et al. (2007) provide evidence that the transactions costs households must undergo to enroll in the program, the level of benefits, and any restrictions that states may impose on Lifeline customers, such as no Caller ID, significantly affect Lifeline participation rates. Our empirical work examines whether transaction costs may be mitigated by state programs that automatically enroll eligible households for Lifeline and Linkup.

[11] Furthermore, as discussed earlier, restructured Lifeline subsidies substantially *lowered* the monthly price to low-income households. Thus, new subscribers as a result of the policy would likely be Lifeline customers.

[12] Nearly every Lifeline usage-based plan includes a small allowance of local calls that the customer can complete for no additional charge.

[13] Vermont also requires a usage-based plan, but is not included in our dataset. Washington requires Lifeline customers subscribe to a flat-rated plan, while Maryland, Arkansas and West Virginia require Lifeline customers subscribe to a usage-based plan. In each of these states non-Lifeline customers may subscribe to either flat-rate or usage-based plans.

[14] *Lifeline*X is the minimum basic monthly charge plus usage charges across all available plans assuming the customer completes X three-minute local calls. The monthly charge component equals the non-Lifeline monthly charge, including the federal subscriber line charge (SLC), less the total Lifeline discount; *Lifeline*X includes extended area of service surcharges when such surcharges are non-optional.

variables *White, Black, Native* (Native Americans)*, Asian* and *Other* (other non-white populations); these census variables are equal to the percentage of low-income populations belonging to the respective group. These are the $\left(X_{1l},...X_{Ml}\right)$'s from the prior section.

An important characteristic of the service is the number of people within a customer's local calling area (LCA). Customers with flat-rate service can make an unlimited number of calls to customers located within their LCA. When subscribing to a usage-based plan, the rates for local calls are lower than charges for calls outside the customer's LCA. The independent variable *LCA* is equal to the number of households within a customer's local calling area.[15] We expect a positive relationship between *LCA* and *Penetration* holding other factors constant.

We also consider two additional service characteristic variables. Three states had programs that automatically enroll eligible households for Lifeline and Linkup.[16] Such programs lower the transaction cost of obtaining subsidized service. The dummy variable *Autoenroll* is equal to one if the state had such an automatic enrollment program. As a robustness check, we also consider the variable *Autoenroll2,* which includes three additional states that adopted programs that reduce the transaction costs associated with enrolling in the Lifeline and Linkup programs.[17] The second additional service characteristic is the level of interstate toll rates. Because customers make both local and long-distance calls, the price of long-distance calls affects subscriber decisions, as emphasized by Hausman, Tardiff, and Belinfante (1993). Intrastate access charges are the fees that local exchange carriers charge long-distance companies for non-local intrastate calls. Thus they capture the effect that the prices of calls outside a customer's LCA have on penetration. The variable *Access*, which is equal to the access charge

---

[15] In wire centers that serve more than one locality, the household-weighted average LCA is used for the wire center. *LCA* is constructed from tariffs, census data, Telecordia (2000), and Claritas (2003).

[16] The FCC (2003) reports that three states – MA, NY and ND –have automatic enrollment programs. In Massachusetts, households that qualify for the low-income heating assistance program (LIHEAP) are allowed to have the LIHEAP-administrating office contact Verizon and enroll them in the Lifeline program. The New York Department of Family Assistance (NYDFA) automatically enrolls a household in the Lifeline and Linkup program when it enrolls in a NYDFA program. The North Dakota Department of Human Services sends certificates to households that allow them to enroll in Lifeline and Linkup programs when they are determined eligible for a program that qualifies them as eligible for Lifeline and Linkup. Information from Center for Media Education/Center for Policy Alternatives (1999) and local tariffs were used to verify that these programs were in place on January 1, 2000.

[17] *Autoenroll2* is set equal to one for California, Maine and Minnesota, in addition to the *Autoenroll* states. California allows customers to self-certify that they meet the eligibility standards. According to the Maine PUC (2000), the Maine Telephone Education Fund, sent a mailing to people eligible for the Lifeline and Linkup programs. Jackson, Baker and Wilden (2002) report that the Minnesota Department of Human Resources certifies eligibility for the program and informs local telephone companies when their subscribers are found to be eligible for the program.

for a four-minute intrastate long-distance call, is expected to have a negative coefficient.[18] We control for only intrastate access charges because interstate long-distance prices in states do not vary with each state's interstate access charge. Section 254(g) of the Act forbids interstate carriers from charging different rates in different states.

We also control for the median income and population density, thus allowing that telephone service is more (or less) valuable in higher income and less rural/more urban communities. The variable *Median Income* is equal to the median income (in $1000s) of households served by a the wire center, the variable *Rural* is the percent of wire-center households living in rural areas, and the variable *MSA* is the percent of wire-center households living in a metropolitan statistical area.

Endogenous Variables and Instruments

We consider the possible endogeneity of three of the explanatory variables: *Lifeline50*, *Lurhook*, and *Autoenroll.* As explained earlier, we are particularly concerned about the possible endogeneity of Lifeline rates because the magnitude of the increases in Lifeline subsidies after 1996 varied significantly across states. Endogeneity could arise if state regulators set these subsidies based on ξ_l, i.e. unobserved (to the econometrician) service characteristics or characteristics of the low-income population.

In considering the possible endogeneity of *Lifeline50*, it is useful to suppose that regulators choose an appropriate subsidy for low-income households, and then subtract this from the normal monthly price:

$$Lifeline50_l = Monthly50_l - Subsidy50_l$$

where *Monthly50* is the normal minimum monthly expenditure for 50 calls, and *Subsidy50* is the discount offered to Lifeline-eligible low-income households. Since the *Subsidy50* component is directed specifically at low-income households, and since these subsidies were significantly increased in 1996, it seems quite plausible that $Subsidy50_l$ is correlated with ξ_l, the unobserved

[18] The access charge used includes per minute originating and terminating carrier common lines charges (CCLC), plus switched access, transitional and call-set up charges, plus any charges for state universal service programs. We assume a four-minute call. In New Jersey, Maryland, Virginia and West Virginia, the CCLC is determined by a long-distance carrier's share of total intra-state long-distance minutes. In these states, the state commission determines the total amount of money to be recovered through the CCLC and charges carriers on a retroactive basis. We estimate the CCLC in these states using ARMIS (FCC, 2000b) and tariff data.

component of demand for low-income households in 2000. For example, correlation might arise from political pressure for higher Lifeline subsidies in areas with lower low-income penetration rates. On the other hand, the *Monthly50* component of *Lifeline50* is a price paid by all subscribing households in an area. Presumably, regulators primarily take *non-low-income* households into account when setting *Monthly50*, since *Subsidy50* can always be adjusted to generate a desired price for low-income households. In addition, *Monthly50* prices tend to change fairly slowly over time, so there may be an important historical component to these prices. Hence we believe a-priori that it is more likely *Subsidy50* is correlated with ξ_l than *Monthly50*. Our empirical work, however, considers both possibilities.[19]

Our other two potential endogenous variables are *Lurhook* and *Autoenroll.* Our a-priori view is that, even though these variables are also targeted towards low-income households, they are less likely to be endogenous than *Lifeline50*. The Federal-State Joint Board on Universal Service (FCC, 2003) recognizes that implementing automatic enrollment procedures imposes additional administrative burdens and costs, suggesting that this policy decision was determined primarily by infrastructure considerations, i.e. whether state computer systems were up to the task. Regarding the possible endogeneity of *Lurhook*, an important issue is the way in which the Federal Government funds low-income subsidy programs, and the resulting incentives for states. As discussed in the introduction, the Lifeline and Linkup programs differ in the extent to which the Federal Government provides matching incentives. In the Lifeline program the federal subsidy increases with the amount of state subsidy, i.e. the state subsidy is "matched." In contrast, in the Linkup program the federal subsidy is fixed at 50% of the customary rate (up to $30). Thus any state subsidization of the Linkup rate is not matched by the federal government. Presumably in response to these strong economic incentives, 36 (81.8%) of the 44 states in our sample provide additional Lifeline subsidies, while only 12 (27.3%) provide additional Linkup subsidies.[20] Summing up, it appears that *Lurhook* is determined for the most part by customary

[19] We emphasize that just because *Monthly50* is set for non-low-income households does not guarantee that it is uncorrelated with ξ_l for low-income households. For example, demand shocks may be correlated across low-income and non-low-income households, and *Monthly50* might be set in response to non-low-income demand shocks. Or, *Monthly50* might be set in response to unobserved product characteristics that affect both non-low-income and low-income demand.

[20] There may be political reasons why states primarily subsidized Lifeline. While Linkup subsidies target a small group of eligible households who have not adopted yet or have just moved, Lifeline subsidies benefit all low-income households. Hence Lifeline subsidies may be politically more feasible.

rates and a fixed federal subsidy percentage.[21] Consequently, it seems plausible that *Lurhook* is exogenous with respect to unobserved state-level variation in low-income demand conditions. Furthermore, *Autoenroll* appears to be determined primarily by plausibly exogenous technological constraints. Therefore, we are less concerned with the possible endogeneity of *Lurhook* and/or *Autoenroll* than with *Lifeline50*. Our empirical work, however, considers possible endogeneity of all three variables.

When allowing for the possible endogeneity of *Lifeline50*, *Lurhook*, and *Autoenroll*, we need valid instruments for identification. The instruments must be variables that exogenously shift the relevant endogenous variable, but do not directly shift low-income demand and are uncorrelated with demand residuals.[22] More intuitively, we want "cost-shifters" that affect the subsidized rates (and/or *Autoenroll*) but are unrelated to low-income demand. Our primary instruments are *State Rural, Competition, Elect PUC,* and *Democrat PUC*. *State Rural* is the percent of rural households in the state. This is interpreted as a proxy for the telephone company's average cost of service in the state because the average cost of service generally decreases with population density.[23] Higher cost is expected to increase prices because state regulators are required to set rates that recover carriers' costs of service.[24]

The variable *Competition* (FCC, 1996) measures whether a state had allowed competitive entry in 1995, before passage of the Act, and whether competitors had begun providing local switched services in a state by 1995. Knittel (2004) finds that the introduction of competition before the Act reduced the amount of cross-subsidization present in local telephone markets. Specifically, he shows that residential prices were higher and business prices were lower in states with active competition.

---

[21] Like *Monthly50*, the customary connection charge is set for all households, not specifically for low-income households. Also like *Monthly50*, this customary rate typically does not vary much over time.

[22] In particular, we maintain that the instruments are uncorrelated with unobserved service characteristics. While this may be a strong assumption, it is very common in the differentiated products demand literature (e.g. Berry, Levinsohn, and Pakes (1995)).

[23] Like Rosston, Savage and Wimmer (2008), we also considered the BOCs' average forward-looking cost of service constructed from the FCC (2000a) Hybrid Cost Proxy Model (HCPM) as an alternative proxy, but discovered that *State Rural* had more explanatory power. State regulation of rates is generally based on historical rather than forward-looking cost.

[24] Recall that we include the percent of households living in rural areas in the wire center as an explanatory variable (*Rural*). Hence, we do allow the level of ruralness in a location to affect telephone demand in that location. The instrument *State Rural* capitalizes on the fact that subsidies are primarily set at the state level. This generates across-state variation in the subsidy conditional on the level of ruralness in a location. We also considered the state poverty rate as an instrument, but found that it did not change our estimates very much.

Democrat PUC and *Elect PUC* describe the state's Public Utility Commission (PUC) and come from NARUC (2000). These commissions played a major role in the determination of states' Lifeline and Linkup subsidies. *Democrat PUC* equals the percentage of a state's PUC commissioners affiliated with the Democratic party, and *Elect PUC* is a dummy variable indicating if state public utility commissioners are elected rather than appointed. Democrats might be more inclined to provide larger subsidies for the poor and elected officials may be more sensitive to the contributions of regulated utilities and set higher residential rates (Rosston, Savage and Wimmer, 2008).

Finally, recalling our decomposition of *Lifeline50, Monthly50* is an additional potential instrument in specifications for which only *Subsidy50* is endogenous. Intuitively, using *Monthly50* as an instrument exploits the part of the variation in *Lifeline50* that is not directed at low-income households as exogenous variation. This should be a particularly strong instrument, since *Monthly50* is mechanically related to *Lifeline50*. In cases where we use *Monthly50* as an additional instrument, we drop the instrument *State Rural* because our arguments above hypothesize that *State Rural* affects prices primarily through the normal rate. Therefore, if we use *Monthly50* as an instrument, *State Rural* is theoretically redundant.

Table 1 provides summary statistics for the full sample of 7,117 wire centers as well as for a restricted sample of 6,596 observations that drops the 521 wire centers with 50 or fewer poor households. We use this restricted sample in our first approach to estimation, because, as explained further below, our first approach is not strictly valid when there is sampling error, i.e. differences between the expected penetration rate in a location and the actual penetration rate in a location. These differences will tend to be smaller in locations with larger populations.

Estimation: First Approach

Following equation (1), our first approach to estimation uses the econometric model:

$$\begin{aligned}\ln Penetration_l &= \theta_0 + \theta_1 Lifeline50_l + \theta_2 Linkup_l + \theta_3 Autoeroll_l \\ &+ \theta_4 \ln LCA_l + \theta_5 Median\ Income_l + \theta_6 Rural_l + \theta_7 MSA_l + \\ &\ln(White_l + e^{\theta_8} Black_l + e^{\theta_9} Native_l + e^{\theta_{10}} Asian_l + e^{\theta_{11}} Other_l) + \xi_l\end{aligned} \tag{3}$$

This assumes that ψ is linear (i.e. $-\lambda\psi(R) = \theta_1 Lifeline50_l + \theta_2 Linkup_l$) and that $\lambda V = \theta_0 + \theta_3 Autoeroll_l + \theta_4 \ln LCA_l + \theta_5 Median\ Income_l + \theta_6 Rural_l + \theta_7 MSA_l + \xi_l$, where ξ_l is the econometric residual, i.e. unobserved service characteristics or demand shocks. θ_8 through θ_{11} correspond to the $\lambda\mu_g$ for the various demographic groups, and $\lambda\mu_{White}$ has been normalized to 1 since it is not separately identified from the constant term θ_0.

Note that we have replaced the expected penetration rate $\overline{S}$ in equation (1) with the observed penetration rate, *Penetration*. This ignores sampling error, i.e. it ignores the fact that the observed penetration rate will vary around $\overline{S}_l$ because the number of poor households is finite. While one can loosely interpret this sampling error being subsumed into the residual ξ_l, this is not a strictly valid interpretation.[25] This is one limitation of our first estimation approach, and is the reason why we use the restricted sample in our first approach. Another limitation of this first approach is that the estimating equation is only strictly valid if no demographic group in any location adopts with probability one. This amounts to assuming an upper bound on the support of the distribution of ξ_l. On the other hand, the virtue of the first approach is that it enables estimation by simple nonlinear IV/GMM techniques, which require neither a full distributional assumption on ξ_l nor a precise specification of the data generating process for potentially endogenous prices.

Because of the non-linearity of the model and the possible endogeneity of explanatory variables, we estimate the model using generalized method of moments (GMM). Our basic moment assumption for estimation is:

$$E\left[\xi_l \otimes Z_l\right] = 0 \tag{4}$$

i.e. the residuals ξ_l are uncorrelated with instruments Z_l. The composition of Z_l varies across specifications of the model depending on the specific exogeneity assumptions. As discussed above, *LCA*, *Median Income*, *State Rural*, *MSA* and the demographic variables are always treated as exogenous, so they always enter Z_l. *Lifeline50*, *Linkup*, and *Autoenroll* enter Z_l when they are treated as exogenous. When any of these variables are treated as endogenous, they are removed from Z_l and replaced by the instruments *State Rural*, *Competition*, *PUC Elect*, and *Democrat PUC*. In specifications where only the subsidy component of *Lifeline50* is treated as endogenous (i.e. the *Monthly50* component of *Lifeline50* is assumed exogenous), the instrument *Monthly50* replaces *State Rural*.

[25] If the dependent variable were $\overline{S}$ rather than $\ln\left(\overline{S}\right)$ it would be a strictly valid interpretation.

Given any arbitrary parameter vector, the implied residuals, $\xi_l(\theta)$, can be computed using (3). At the true parameter vector θ *, the implied $\xi_l(\theta^*)$'s equal the true residuals; at other parameter vectors this is assumed not to be the case. Thus, estimation proceeds by considering:

$$G(\theta) \equiv E\left[\xi_l(\theta) \otimes Z_l\right] \approx \frac{1}{N}\sum_l \xi_l(\theta) \otimes Z_l \equiv \frac{1}{N}\sum_l g_l(\theta) \equiv G_N(\theta)$$

Given the orthogonality assumption on the true residuals given by (4), $G(\theta)$ (and, asymptotically, $G_N(\theta)$) equals 0 when evaluated at θ *; at other parameter vectors, this is assumed not to be the case. Hence, a consistent estimator is obtained by searching for the θ that makes $G_N(\theta)$ "as close as possible" to zero. Formally, this is done by minimizing a quadratic form in $G_N(\theta)$, i.e.

$$G_N(\theta)' A G_N(\theta) \tag{5}$$

where A is a full rank weight matrix that only affects efficiency, not consistency. The weight matrix A that minimizes the variance of the resulting estimate is A=Var($G_N(\theta)$)$^{-1}$. We use a standard two-step procedure to approximate this optimal weight matrix A.

There are two additional econometric considerations we address in our estimation procedure. The first concerns the fact that our observations represent geographic areas. Thus, one might expect the unobservables ξ_l to be correlated across nearby wire centers, e.g. if some aspects of unobservable service characteristics ξ_l are determined at the state level. While this does not affect the consistency of our GMM estimators, it does impact their standard errors. To address this, we allow for geographic clustering at the state level (i.e. we allow for correlations in the residuals across wire centers in the same state) in computing these standard errors.

To allow for such "clustering," we use a robust variance estimator (e.g. Moulton, 1990) to compute standard errors. Formally, the variance of the GMM estimator that minimizes (5) is given by

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

where

$$V = Var(G_N(\theta)) \qquad \text{and} \qquad \Gamma = \frac{\partial G(\theta)}{\partial \theta'} \approx \frac{\partial G_N(\theta)}{\partial \theta'}$$

With independent observations, V could be estimated consistently with

$$\hat{V} = \frac{1}{N^2}\sum_l g_l(\theta) g_l(\theta)'$$

evaluated at the estimated θ. However, if there are correlations between the ξ_l's, this is not a consistent estimate of V. To address such clustering at the state level, we can also write V as:

$$V = Var\left(\frac{1}{N}\sum_{s}\sum_{l\in s} g_l(\theta)\right) = Var\left(\frac{1}{N}\sum_{s}\Phi_s(\theta)\right)$$

where s indexes states, $l \in s$ indexes wire centers in state s, and $\Phi_s(\theta) = \sum_{l\in s} g_l(\theta)$.

Under the assumption that the residuals are independent *across* states, the Φ_s's are independent across s and V is consistently estimated with

$$\hat{V} = \frac{1}{N^2}\sum_{s}\Phi_s(\theta)\Phi_s(\theta)'$$

evaluated at the estimated θ. Note that this $\hat{V}$ is consistent regardless of the patterns of the residual correlation within a state, or differences in the pattern of correlations across states. For example, there might be higher correlation between ξ_l the closer are two locations in the same state, and the ξ_l might be more highly correlated in some states than others.[26] $\hat{V}$ is also a consistent estimator of V whether or not there is heteroskedasticity in the residuals.

The second non-standard issue concerns the fact that wire center penetration data are aggregated. Since the number of households, N_l, varies across wire centers, we expect the aggregation to generate heteroskedastic residuals. While such heteroskedasticity does not affect the consistency of our estimates, it is possible to gain efficiency by addressing the heteroskedasticity appropriately. We do so by introducing weights into the estimation procedure. We first estimate the model ignoring heteroskedasticity, and then linearly regress the squared estimated residuals, $\hat{\xi}_l^2$, on functions of N_l to estimate how the variance of the residuals depends on population size. We use these regression results to construct weights w_l equal to the inverse of the square root of the predicted variance for each location. We then re-estimate the model, using weighted residuals $w_l\xi_l$. These weights equalize the variance of the weighted residuals across observations, and therefore are optimal by construction.

Table 2 presents our first set of estimates. The different columns correspond to the different endogeneity specifications discussed earlier. Column 1 is the ALL EXOGENOUS specification. While perhaps unrealistic, this provides a point of comparison for the less

[26] On the other hand, this rules out correlation between wire centers that are nearby geographically, but in different states. This assumption is more reasonable if the residuals represent unobserved service characteristics that are set by states.

restrictive estimators. Columns 2 through 9 allow for increasingly less restrictive exogeneity assumptions, starting with Column 2, in which only the subsidy component (*Subsidy50*) of *Lifeline50* is treated as endogenous, and ending with column 9, where *Lifeline50, Linkup,* and *Auto* are all treated as endogenous.

First, note that the coefficients on the demographic groups, *Rural, ln(LCA), Median Income,* and *MSA* change very little across the various specifications. Given the normalization of the coefficient on *White*, the coefficients on the other demographics measure the strength of demand of these demographics relative to the white population, i.e. $\theta_g = 0$ means that group *g* has the same demand as white households. Thus significant negative coefficients on *Black*, *Native*, and *Other* indicate that, all else equal, these groups have lower penetration rates than whites. In contrast, the positive coefficients on *Asian* indicate that Asian households have higher penetration rates. As expected, the coefficient on *ln*(*LCA*) is positive and significant across all the specifications. The magnitude of the effect seems reasonable – a doubling of the local calling area increases penetration rates by almost 1%. Also as expected, the coefficient on *Median Income* is positive and statistically significant. The coefficient on *Rural* is negative and generally significant, and the coefficient on *MSA* is significant and positive – suggesting unobserved service characteristics are better (or that telephone demand is stronger) in more urban areas.

We are concerned primarily with the estimated coefficients on *Lifeline50*, *Linkup*, and *Autoenroll*. In addition to reporting the coefficients themselves, we compute some interesting functions of the coefficients. First, we report the price elasticities implied by the coefficients on *Lifeline50* and *Linkup* (*Elasticity50* and *ElasticityLU*). These elasticities are evaluated at the sample median.[27] Second, while both price variables are measured in dollars, *Linkup* is a one-time connection fee while *Lifeline50* is a recurring monthly fee. Hence we can use the relationship between the two coefficients to approximate the rate at which these households discount the future. Formally, the monthly discount rate implied by the price coefficients is computed as $Discount = \theta_2 / \theta_1$.

The ALL EXOGENOUS specification in Column 1 establishes a baseline. The price coefficients and elasticities are negative but very small, as expected from previous research on

[27] Given our log-linear penetration equation, the reported price elasticity is equal to the estimated coefficient times the median price in the sample.

telephone demand. The elasticities with respect to *Lifeline50* and *Linkup* are -0.01025 and -0.00611 respectively, and only the former is statistically significant. The coefficient on *Autoenroll* in Column 1 is a statistically significant 0.022, suggesting that automatic enrollment policy increases low-income penetration rates by 2.2%. Since penetration rates are already quite high (averaging about 92% in our population), an automatic enrollment policy reduces the number of households without service by almost 30%, which makes sense given the low take-up of Lifeline and Linkup programs in states lacking automatic enrollment policies.[28] Given that prices enter the penetration equation linearly, the coefficient on *Autoenroll* divided by θ_2 (the coefficient on Linkup) can be interpreted as the reduction in the fixed transaction cost of initiating Lifeline service resulting from an automatic enrollment policy. The implicit reduction is equal to about \$45, suggesting that an automatic enrollment policy has substantial value to consumers. In our preferred specifications (columns 2 and 3) the effect falls to around \$36, because θ_2 is higher in these specifications. Finally, the discount rate is almost 25% per month, although it is estimated poorly. The high discount rate is not unreasonable for poor households facing credit constraints.[29]

Column 2 treats the subsidy portion of *Lifeline50* (i.e. *Subsidy50*) as endogenous, using *Competition, Elect PUC, Democrat PUC*, and *Monthly50* as instruments. As argued in the prior section, our a-priori belief is that *Subsidy50* is the price component that is most likely to be endogenous. A first observation concerns the F-statistic from the corresponding first stage regressions, reported at the bottom of the table.[30] At 14.16, the F-statistic easily rejects the null hypothesis that the instruments are not predictive of the endogeneous variable, suggesting that the instruments are strong enough for meaningful inference. Moving to the coefficient estimates, the coefficient on *Autoenroll* is similar to that in Column 1, while the price coefficients (and implied elasticities) are considerably larger in absolute value. The *Lifeline50* coefficient/elasticity increases by more than 50%, and the *Linkup* coefficient/elasticity increases by about 40%. In addition, the *Linkup* coefficient becomes significant at the 90% level. The

[28] The FCC (2003) estimates that about 33 percent of the low-income households that were eligible for the Lifeline and Linkup programs in 2000 participated in the programs. The Center for Media Education/Center for Policy Alternatives (1999) also reports low participation rates.

[29] See, for example, Adams, Einav, and Levin (2009), who study low-income populations in the context of used car loans. If there is heterogeneity in these discount rates across the population, we are most likely identifying the high end of the range, because the "marginal" households determining our estimated coefficients are most likely the poorest of the poor.

[30] These F-statistics were computed in STATA and are robust to clustering by state.

increase in the *Lifeline50* coefficient between Columns 1 and 2 is intuitive. If the *Subsidy50* component of *Lifeline50* is endogenous and correlated with the demand residual, we would expect the correlation to be positive, i.e. we would expect states with lower penetration rates to more heavily subsidize (i.e. set lower prices).[31] This endogeneity would generally cause the *Lifeline50* coefficient in Column 1 to be biased towards zero.[32] The estimated discount rate in Column 2 decreases slightly from Column 1, but it is still not statistically significant.

We can use the estimates in Columns 1 and 2 to formally test whether the *Subsidy50* component of *Lifeline50* is endogenous. The Hausman test statistic[33] for this comparison is 20.05, which is significant at the 90% confidence level. This suggests that *Subsidy50* is in fact correlated with the demand unobservable, and implies that the estimates in Column 2 are preferable to those in Column 1. Since there are more excluded instruments (4) than endogenous variables (1) in Column 2, the model is overidentified and we can also run a specification test. A J-test of these over-identifying restrictions (test statistic = 1.36) does not reject the validity of the moment conditions.

Column 3 relaxes our exogeneity assumptions a bit further, allowing the entirety of *Lifeline50* to be endogenous. Relative to the estimates in Column 2, we drop *Monthly50* as an instrument (since it is now being considered endogenous), and add *State Rural* as an instrument (since it is assumed to be an exogenous determinant of *Monthly50*). Interestingly, the results do not change much from Column 2, though the *Lifeline50* coefficient increases slightly. Again, the instruments appear strong enough (F-stat = 16.80), the specification passes the overidentifying test (J-test statistic = 1.26), and the Hausman test rejects the ALL EXOGENOUS specification (Hausman test statistic = 22.86). We also performed a Hausman test comparing Column 2 to Column 3 to test the null hypothesis that *Monthly50* is exogenous. We are unable to reject the null, though this might be because the test lacks power.

[31] There are other reasons for a positive correlation, for example, service with "better" unobserved characteristics might be more costly to provide, and hence be priced higher.

[32] As discussed above, prior studies have found low elasticities of demand. While these studies are not directly comparable due to differences in the time period of the data and the techniques, endogeneity may have played a part in the low estimates.

[33] Since we are in a nonlinear framework, our ALL EXOGENOUS specification does not necessarily yield an efficient estimator. Hence, we cannot use the standard Hausman formula to derive the covariance between each set of estimates (e.g. the correlations between the estimates in column 1 and the estimates in column 2). To derive these covariances, we estimate both specifications simultaneously (using moments from both specifications with two sets of parameters - one entering each set of moments). As a result, the standard GMM variance formula gives us the covariances between the parameters of the two specifications that we need for the Hausman test.

Columns 4 through 9 further relax our exogeneity assumptions. Columns 4-7 start from the specifications in Columns 2 and 3 and alternatively allow *Linkup* and *Autoenroll* to be endogenous. Columns 8 and 9 start from the specifications in Columns 2 and 3 and allow both *Linkup* and *Autoenroll* to be endogenous. An unfortunate result in all these specifications is that our excluded instruments are very weak for these additional potentially endogenous variables. The F-stats for *Linkup* and *Autoenroll* (when they are treated as endogenous) are not significant by conventional standards, let alone the standards of the weak instrument literature. The weakness of the instruments manifests itself in the large standard errors of the *Linkup* and *Autoenroll* coefficients in these specifications.

While there is no statistical evidence suggesting that *Linkup* and *Autoenroll* are endogenous, this may be due to the weakness of the instruments and resulting high standard errors.[34] However, several arguments give credence to our choice of Columns 2 and 3 as preferred estimates: 1) our arguments from the prior section that *Lifeline50* (or its component *Subsidy50*) is the most likely of the policy variables to be endogenous; and 2) the fact that the coefficient on *Lifeline50* is relatively stable across specifications.

Table 3.a considers a first set of simple robustness checks on the model. All are perturbations of Column 2 in Table 2, i.e. the model where only *Subsidy50* is considered endogenous.[35] Column 1 uses *Lifeline100* as the monthly price instead of *Lifeline50*. As described earlier, this variable measures the minimum monthly expenditure of Lifeline customers making 100, rather than 50, local calls. Column 2 uses our alternative definition of automatic enrollment, *Autoenroll2*, as described in the Data section. Column 3 drops the *Autoenroll* variable altogether, and Column 4 adds the *Access* variable, a measure of intrastate long distance calling rates. Finally, column 5 estimates the logarithmic model described in the Model section where $\psi(R) = \ln R$. This leads to the econometric model:

$$\begin{aligned}
&\ln Penetration_l = \theta_0 + \theta_1 \ln(\theta_2 Lifeline50_l + Linkup_l) + \theta_3 Autoeroll_l \\
&+\theta_4 \ln LCA_l + \theta_5 \ln Median\ Income_l + \theta_6 Rural_l + \theta_7 \ln MSA_l + \\
&\ln(White_l + e^{\theta_8} Black_l + e^{\theta_9} Native_l + e^{\theta_{10}} Asian_l + e^{\theta_{11}} Other_l) + \xi_l
\end{aligned}$$

[34] We do not report the results in the paper, but no Hausman test comparing Columns 4-9 to their restricted analogues in Column 2 or 3 rejects the null hypothesis that either *Linkup* or *Autoenroll* is exogenous. Of course, given the weakness of the instruments for *Linkup* and *Autoenroll*, these tests are probably not very powerful.

[35] Similar robustness checks for the Lifeline50 ENDOGENOUS specification are reported in Table 3.b. The results are qualitatively similar.

where $price = \theta_2 Lifeline50 + Linkup$ is the capitalized price of service and where $1/\theta_2$ is the implied monthly discount rate. Given the alternative specification, it is hard to compare the estimated price coefficients in Column 5 to the other models. However, the implied price elasticities are comparable.

The estimates in Table 3.a suggest that our results are quite robust. The Lifeline and Linkup elasticities do not move by more than about 35%, and generally stay significant (the only exception being the *Linkup* coefficient, which becomes insignificant in Columns 2 and 3). Similarly, the *Autoenroll* coefficient is very stable across the specifications, and remains highly significant. Interestingly, the coefficient on *Access* is not close to being significant (and the wrong sign). The failure to find a significant negative effect of intrastate long distance prices on penetration is contrary to Hausman, Tardiff, and.Belinfante (1993), but may be explained by the fact that intrastate access prices were lower and had limited variation in our 2000 cross section.[36] The high standard error of our estimate certainly does not rule out a quantitatively significant negative effect of intrastate long distance prices on penetration.

Table 4.a presents four additional robustness checks, all having to do with how we restrict our sample.[37] Column 1 drops California from the sample. Unlike other states, California did not require formal verification of eligibility for Lifeline and Linkup programs. As such, California has an extremely high take-up rate.[38] Given the size of the state, one might be concerned that this peculiarity may be driving some of our results. Columns 2-4 consider alternative ways to address locations with small poor populations. Recall that because our first approach to estimation does not explicitly consider sampling error, our results in Tables 2-3 restrict the sample to locations with more than 50 poor households. In Column 2 of Table 4.a, we alternatively include all locations in the sample, and in Column 3, we restrict the sample to locations with more than 100 poor households. Lastly, Column 4 additionally drops locations where the observed adoption rate of poor households is 100%. Even though this last selection criteria is not econometrically valid (since the selection criteria is based on the dependent

[36] Hausman, Tardiff, and Belinfante (1993) used panel data covering 1984-1988. The panel data set allowed them to exploit variation in interstate as well as intrastate long distance prices. Interstate prices fell substantially during this period due to reduced access prices.

[37] Similar robustness checks for the Lifeline50 Endogenous specification are reported in Table 4.b. Again, results were qualitatively similar.

[38] FCC (2003, Appendix F) estimates that more than 100% of California households eligible for Lifeline received subsidies in 2000.

variable), it still seems like a reasonable check on the robustness of our results to 1) not formally modeling sampling error, and 2) assuming away 100% adoption probabilities. Again, our second estimation approach will explicitly allow for sampling error and 100% adoption probabilities.

The estimates in Table 4.a suggest that our results are quite robust. The implied elasticities move very little, and the *Lifeline50* and *Autoenroll* coefficients remain significant; again the *Linkup* coefficient loses significance in two of the specifications.

In summary, Columns 2 and 3 in Table 2 are our preferred specifications from our first estimation approach. These two specifications control for the possible endogeneity of *Lifeline50*, but treat all other explanatory variables as exogenous. The estimated price elasticities are small but higher than previous studies have found for the entire population, a conclusion that shows controlling for endogeneity matters. The estimated model also shows consumers value larger local calling areas, and that an automatic enrollment policy for Lifeline and Linkup substantially boosts the telephone penetration of low-income households. There are significant demographic differences in the demand for service. These results appear robust to a number of modeling perturbations.

Estimation: Second Approach

Our second approach to estimation addresses sampling error explicitly and allows for the possibility that members of some groups, in some locations, might adopt with probability one. To do this, we need to additionally assume a fully parameterized model of the process determining Lifeline prices as well as the full joint distribution of that process and the unobservables in the adoption model.

Consider the theoretical framework in Section 2, with data on N_l – the number of low-income households in location l, X_{gl} – the share of those households in demographic group g in location l, $Penetration_l$ – the overall adoption rate in location l, and the explanatory variables (the variables on the right hand side of (3)) across locations.

Our underlying theory of household demand implies that the distribution of $Penetration_l$ conditional on the other observables *and* the "unobserved service characteristic" ξ_l is given by:

$$
\begin{aligned}
&Penetration_l = \sum_g X_{gl} S_{gl}; \\
&S_{gl} X_{gl} N_l \sim Binomial\left(X_{gl} N_l, \overline{S}_{gl}\right); \\
&\ln \overline{S}_{gl} = \min\{0, C_g + \theta_1 Lifeline50_l + \theta_2 Linkup_l + \theta_3 Autoeroll_l \\
&+\theta_4 \ln LCA_l + \theta_5 \ln Median\ Income_l + \theta_6 Rural_l + \theta_7 MSA_l + \xi_l\}
\end{aligned}
\tag{7}
$$

$X_{gl} N_l$ is the number of individuals in group *g* at location *l*. The C_g's are group specific constants which are related to the coefficients on the population shares in the first model. More specifically, $C_{white} = \theta_0$, $C_{black} = \theta_0 + \theta_8$, *etcetera*. This model is equivalent to the previous model except that equation (3) assumes 1) the min function in (7) never binds at 0, and 2) strictly speaking, $Penetration_l = \sum_g X_{gl} \overline{S}_{gl}$.[39] In contrast, the current model explicitly models the variance of observed $Penetration_l$ around $\sum_g X_{gl} \overline{S}_{gl}$ due to the binomial sampling process.

We make the following assumption on the model regarding the determination of Lifeline prices in each location:

$$Lifeline50_l = \theta_{12}' Z_l + \eta_l$$

where Z_l is a set of exogenous observables, and η_l is an unobservable. Substituting this "first stage" equation into the equation determining $\overline{S}_{gl}$ results in a set of two "reduced form" equations:

$$
\begin{aligned}
&Lifeline50_l = \theta_{12}' Z_l + \eta_l \\
&\ln \overline{S}_{gl} = \min\{0, C_g + \theta_1\left(\theta_{12}' Z_l + \eta_l\right) + \theta_2 Linkup_l + \theta_3 Autoeroll_l \\
&+\theta_4 \ln LCA_l + \theta_5 \ln Income_l + \theta_6 Rural_l + \theta_7 \ln MSA_l + \xi_l\}
\end{aligned}
\tag{8}
$$

While $\overline{S}_{gl}$ is not directly observed, it defines the distribution of $Penetration_l$ in equation (7).

In light of our previous discussion and results from the first approach, we focus on the possibility that *Lifeline50* is endogenous, while maintaining the assumption that *Lurhook* and *Autoenroll* are exogenous. We consider three different assumptions on the endogeneity of *Lifeline50*, which imply different Z_l. First, the ALL EXOGENOUS model assumes that all the

[39] While the error term in the first approach could be viewed as including sampling error, this potentially is a source of specification bias because the sampling error in the dependent variable is inside the logarithm.

observables determining $\overline{S}_{gl}$ in equation (7) are in Z_l. Since this includes *Lifeline50*, this means that in this case, the first reduced form equation is simply *Lifeline50* = *Lifeline50* and $\eta = 0$. Second, we consider a specification where *Subsidy50* is treated as endogenous. In this case, *Lifeline50* is dropped from Z_l and replaced with *Competition, Elect PUC, Democrat PUC*, and *Monthly50*. Third, we treat the entire *Lifeline50* price as endogenous. In this case we then replace *Monthly50* with *State Rural.* The rationale for these "instruments" is the same as for our first estimation approach. We complete the model by assuming that the unobservables (η_l, ξ_l) are distributed normally with mean zero and variance matrix Σ. These unobservables are also assumed independent of the binomial process. In the ALL EXOGENOUS specification, the covariance element of Σ is implicitly 0 (since $\eta = 0$); in the other two specifications this covariance is estimated.

We estimate the model using GMM.[40] Specifically, we use the moments

$$E\begin{bmatrix} \left(Penetration_l - E\left[Penetration_l \mid Z_l, N_l; \theta\right]\right) \otimes Z_l \\ \left(Lifeline50_l - E\left[Lifeline50_l \mid Z_l, N_l; \theta\right]\right) \otimes Z_l \\ \left(Penetration_l^2 - E\left[Penetration_l^2 \mid Z_l, N_l; \theta\right]\right) \\ \left(Lifeline50_l^2 - E\left[Lifeline50_l^2 \mid Z_l, N_l; \theta\right]\right) \\ \left(Penetration_l Lifeline50_l - E\left[Penetration_l Lifeline50_l \mid Z_l, N_l; \theta\right]\right) \end{bmatrix} \tag{8}$$

which by our assumptions equal zero when evaluated at the true parameter vector θ^*. The inner expectations in these moment conditions are over both the binomial errors for each group and the bivariate normal errors (η_l, ξ_l). The outer expectation is over the distribution of the observed data $(Penetration_l, Lifeline50_l, Z_l, N_l)$. The distributional assumptions and the exponential form of our model allow analytic computation of the inner expectations, as detailed in Appendix A.

Regarding the specific moments, in the ALL EXOGENOUS specification (where Z_l contains *Lifeline50*), the 2nd and 4th sets of moments can be trivially set to exactly zero by appropriate choice of θ_{12} (i.e. the coefficient on *Lifeline50* equals 1 and all other coefficients equal 0), so these moments can simply be dropped. The last moment can also be dropped, since

[40] MLE is an alternative approach, but is difficult to implement because it requires computing convolutions of independent binomial distributions. GMM sacrifices some efficiency (vs. MLE) but gains computational simplicity and feasibility. The GMM approach also very easily allows for state-level clustering, which is important given that the variation in many of our instruments is at the state-level.

in this case it is redundant to the first set of moments. So estimation of the ALL EXOGENOUS specification uses only the 1st and 3rd sets of moments. The other two specifications use all five sets of moments.

The five moment conditions correspond to specific parameters in our model. For example, the first moment conditon requires that the difference between observed $Penetration_l$ and the expected value of $Penetration_l$ (given Z_l, N_l and the true parameters (θ^*)) is uncorrelated with Z_l. This condition should identify the parameters in the $\overline{S}_{gl}$ equation. Analogously, the next moment condition should identify the parameters in the first-stage price equation, and the remaining three moment conditions identify the parameters of the covariance matrix.

We use two-stage GMM with an optimal weighting matrix. We omit the details for brevity, but in calculating standard errors, we again allow for arbitrary correlations in unobservables across locations within a state. Our estimated standard errors are also consistent with heteroskedasticity in the moment conditions generated by N_l varying across locations.

Table 5 contains our estimates, both for the full sample and for the restricted sample that drops locations with 50 or less poor households. The restricted sample is only for comparison purposes. There is no apparent theoretical reason to drop the small locations in the second approach because the second approach deals with sampling error in a completely consistent fashion. Nevertheless, the results suggest conclusions very similar to those from the first approach. The implied price elasticities in the second approach are slightly higher: when only *Subsidy50* is treated as endogenous, the *Lifeline50* median elasticity is 0.02 and the *Linkup* elasticity is 0.001; when *Lifeline50* as a whole is treated as endogeneous these elasticities are nearly double. The *Autoenroll* coefficients also are higher in the second approach. According to the point estimates (with *Lifeline50* endogenous), *Autoenroll* is equivalent to a $42 signup subsidy. The race coefficients also exhibit similar patterns in the two approaches.[41] As in the first approach, the *F-stat*s suggest that the excluded instruments are strong, and the *J-stat* does

[41] Note that estimated standard errors on the Asian coefficients are extremely high. This is because Asians have very high adoption rates, so the derivative of adoption rate w.r.t. to the Asian coefficient is very close to zero at the point estimate. We are not concerned with this for two reasons. First, we are not explicitly interested in the Asian coefficient, and the standard errors of the other coefficients do not appear to be affected. Second, we believe that there is significant small sample bias in the estimates of these standard errors, because the concavity of the objective function changes dramatically as the Asian coefficient decreases (it is very close to zero at the point estimate). Supporting this supposition is the fact that when we additionally run a restricted specification where $C_{Asian} = C_{White}$, a C-test (a difference in J-statistics test) strongly rejects the null hypothesis that $C_{Asian} = C_{White}$.

not reject the model specification. The significant estimated correlation coefficient in column (3) rejects the hypothesis that that *Lifeline50* is exogenous in these models, although the Hausman tests fail to reject. This robustness of the basic results is at least suggestive that the strong underlying assumptions that differ between the approaches do not bias the estimates in any important way.

Policy Experiment

Using the estimates from our preferred specification (Table 2, Column 2, *Subsidy50 ENDOGENOUS*), we evaluate the impact of the Lifeline and Linkup plans on low-income penetration. We use the estimated penetration equation to see what penetration rates for low-income households would have been if Lifeline and/or Linkup programs were absent.[42] The actual penetration rate for low-income households in our sample is 93.95%. Table 6.a presents two different sets of estimates for the effect of Lifeline and Linkup because different states have different automatic enrollment policies. The first column presents the results of simply eliminating Lifeline or Linkup subsidies while allowing *Autoenroll* to have the same positive impact on predicted penetration even though one of the programs to which the automatic enrollment policies apply is absent. The policy experiment is consistent with the interpretation that automatic enrollment policies reduce the transaction cost of subscribing to subsidized telephone service. The second column shows the results of eliminating Lifeline and Linkup when *Autoenroll* = 0 in all states both before and after the elimination of the low-income support programs. This adjustment results in almost the same estimated impacts from each of the two programs. Finally, the third column shows the total effect of eliminating automatic enrollment, Lifeline, and Linkup policies. The discussion below focuses on the first and third columns; the results from the second column are in parentheses where they differ. The bottom part of Table 6.a shows the very similar effects from the second approach to estimation. Table 6.b shows slightly larger effects from the programs when we use *Lifeline50* ENDOGENOUS estimates (column 3 of Table 2.a) instead of *Subsidy50* ENDOGENOUS.

The predicted penetration rates for low-income households with Lifeline and Linkup rates are significantly and substantially higher than the predicted penetration rates without these

[42] In these policy experiments we assume that normal rates remain unchanged. This is a reasonable approximation since 1) low income households are a small component of overall telephone demand, and 2) normal rates are subject to considerable amounts of regulation.

reduced rates. The estimated difference in the penetration rates of poor households is 4.1 (3.8) percentage points. Most of this increase is explained by the incremental effect of Lifeline, i.e. if Lifeline did not exist then penetration would be lower by 2.3 percentage points. Removing Linkup would reduce predicted penetration of telephone service for poor households by 1.6 percentage points.[43]

To get an idea of the effectiveness of Lifeline and Linkup relative to the costs of the programs, we estimated crudely the amount of federal and state funding for Lifeline and Linkup in our sample. A description of the methodology is in Appendix B. We calculate that the annual federal funding for Lifeline and Linkup in our sample was about $115.6 million in 2000. In addition, we calculate that states spent another $38.5 million on these two programs. There are about 5.2 million low-income households in our sample. A 4.1 (3.8) percentage point increase in penetration among low-income households, means that these programs encourage 213,000 (198,000) more low-income households to subscribe to the telephone network. This works out to a cost of $723 ($780) per poor household per year.[44] Furthermore, there may be additional costs associated with automatic enrollment polices, as well as other implementation costs.

Linkup appears to be much more cost effective than Lifeline. Linkup costs less than 8% of the Lifeline program annually, yet has about two-thirds of Lifeline's incremental effect on predicted penetration. Our estimates suggest that regulators might get the same effect on penetration with substantially less money by increasing the Linkup program and reducing the Lifeline program. The Universal Service Administrative Company (2007) reports that in 2006 the Federal government spent $778 million on Lifeline and only about $33 million on Linkup; so there is room to undertake this policy adjustment. One of the reasons Linkup is more cost effective is that by definition it is targeted at poor households who do not have telephone service. In contrast, we estimate that only about 9% of Lifeline expenses in our sample go to households who would not otherwise subscribe to service.[45]

---

[43] These predictions ignore possibly offsetting factors (Hausman, Tardiff, and Belinfante, 1993). Federal low income subsidy programs are funded by taxes on interstate revenues. To the extent that such extra charges are also borne by low-income households, their bills would decrease somewhat, partially offsetting the increase in hookup and monthly charges.

[44] Our estimate of the cost per household does not account for eligible households whose income exceeds the poverty level. As discussed above, we estimate that over 3 million households with incomes above the poverty level are eligible for the Lifeline and Linkup programs.

[45] We calculate this by adding up the additional households that subscribe due to Lifeline in each wire center times the total lifeline subsidy in that wire center and then dividing the total dollars nationwide by the total cost of the Lifeline program. Calculating the corresponding percentage for Linkup is more difficult because the Linkup subsidy

Conclusions

Using data from 7,117 wire centers, we conclude that low-income subsidy programs have increased low-income telephone penetration by 4.1 percentage points. The conclusion is based on estimated price elasticities of demand with respect to subscription and connection charges for poor households of -0.016 and -0.008 respectively. These estimated elasticities are low but nevertheless somewhat higher than previous estimates for all households. The higher estimates are due substantially to bias corrections that account for the possible endogeneity of Lifeline rates in different locations due to different implementations by state regulators. Even with a relatively low price elasticity of demand, the magnitude of Lifeline and Linkup programs are sufficient to reduce substantially the effective prices faced by low-income households so that telephone penetration increases significantly as result of these programs.

Because of low-income households' high discount rates, the Linkup program has a much higher effect on penetration per dollar spent than the Lifeline program. One possible explanation for this is that low-income households may be credit constrained and even with the typical 50% discount initial hookup charges could be daunting if the expected tenure in the residence is short. Furthermore, Linkup subsidies are targeted at households who do not have telephone service.

The bottom line is that Lifeline and Linkup programs in 2000 connected to the telephone network an additional 213,000 of poor households in our sample at an expense of $723 each.

References

Belinfante, A. (2003) "Telephone Penetration by Income by State," Federal Communications Commission: Washington.

Belinfante, A. (2007) "Telephone Penetration by Income by State," Federal Communications Commission: Washington.

Burton, M., J. Macher and J. Mayo. (2007) "Understanding Participation in Social Programs: Why Don't Households Pick up the Lifeline?" *The B.E. Journal of Economic Analysis and Policy* V. 7, 1: pp. 1-26.

is non-recurring unless a household stops and then renews service. A poor household that moves to a new location might take advantage of Linkup and/or Lifeline when renewing telephone service, and some percentage of these households would subscribe to telephone service without the Linkup and Lifeline subsidies. We are not able to estimate this percentage.

Cain, P. and J. P. MacDonald (1991), "Telephone Pricing Structures: The Effects of Universal Service," *Journal of Regulatory Economics*, 3; 293-308.

Center for Media Education/ Center for Policy Alternatives (1999), "Sorry, No Number!" dated October 21, 1999, posted at www.stateaction.com.

Claritas (2003), "Block Group to Wirecenter Cross Reference," electronic file provided by Claritas, Ithaca, NY.

Crandall, R. and L. Waverman (2000), *Who Pays for Universal Service? When Telephone Subsidies Become Transparent*, Washington, DC: Brookings.

Erikson, R., Kaserman, D. and J. Mayo (1998), "Targeted and Untargeted Subsidy Schemes: Evidence from Post-Divestiture Efforts to Promote Universal Service," *Journal of Law and Economics*, 41: 477-502.

Federal Communications Commission (1996), "Common Carrier Competition".

__________ (2000a), "Hybrid Cost Proxy Model" available at http://www.fcc.gov/wcb/tapd/hcpm/welcome.html.

__________ (2000b) Automated Reporting Management Information System (ARMIS) Report, available at http://www.fcc.gov/wcb/armis/

__________ (2000c) Federal-State Joint Board on Universal Service Promoting Deployment and Subscribership in Unserved and Underserved Areas, Including Tribal and Insular Areas CC Docket No. 96-45, Released June 30, 2000.

__________ (2003) Federal-State Joint Board on Universal Service, CC Docket No. 96-45, FCC 03J-2, Released April 2, 2003, Adopted March 27, 2003

Garbacz, C. and H.G. Thompson, Jr. (2002) "Estimating Telephone Demand with State Decennial Census Data from 1970-1990," *Journal of Regulatory Economics* 21:317-329.

Garbacz, C. and H.G. Thompson, Jr. (2003) "Estimating Telephone Demand with State Decennial Census Data from 1970-1990: Update with 2000 Data," *Journal of Regulatory Economics,* 34: 373-378.

Hausman, J. Tardiff, T. and A. Belinfante (1993) "The Effects of the Breakup of AT&T on Telephone Penetration in the United States," *American Economic Review, Papers and Proceedings*, 83: 178-84.

Jackson, A., C. Baker and R. Wilden (2002). *Energy and Telephone Assistance in the States: Public Programs that Help Low-Income Households* AARP Public Policy Institute. Washington, DC.

Maine Public Utilities Commission, Consumer Assistance Division (2000). *1999 Annual Report on Consumer Assistance.* Augusta, ME.

NARUC (National Association of Regulatory Utility Commissioners) (2000) "Membership Directory," Washington D.C.

Office of Congresswoman Doris Matsui (September 24, 2009) "Matsui Introduces Legislation to Help Low Income Americans Subscribe to the Internet". Press Release.

Perl, L. (1984) "Revisions to NERA's Residential Demand for Telephone Service 1983," prepared for The Central Services Organization, Inc. of the Bell Operating Companies, NERA, White Plains, NY April 24, 1984.

Riordan, M., (2002) "Universal Residential Telephone Service," in Cave, Majumdar, and Vogelsang (eds.) *Handbook of Telecommunications Economics, Vo 1*, Amsterdam: North Holland.

Rosston, G., S. Savage, and B. Wimmer (2008) "The Effect of Private Interests on Regulated Retail and Wholesale Prices," *Journal of Law and Economics,* Vol 51, No 3.

Taylor, L., (1994) *Telecommunications Demand in Theory and Practice,* Dordrecht, Netherlands and Boston: Kluwer.

Taylor, L. and K. Kridel (1990) "Residential Demand for Access to the Telephone Network," in A. deFontenay, M. Shugart, and D. Sibley, eds., *Telecommunications Demand Modelling*, Amsterdam: North Holland.

Telcordia (2000) "Local Exchange Routing Guide." New Jersey.

United States Department of Commerce (2000) "Census Bureau, 2000 sf3 Documentation," U.S. Census Bureau, Economics and Statistics Administration.

Universal Service Administrative Company (2007), "2006 Annual Report".

Table 1

	Full Sample (7117 obs.)					Drop 50 Sample (6596 obs.)*				
	Mean	Median	St.D.	Min	Max	Mean	Median	St.D.	Min	Max
Poor households	1170	419	1970	3	25740	1260	485	2019	51	25740
Penetration	.92	.94	.06	.59	1	.92	.93	.06	.59	1
Lifeline50	5.07	5.19	2.43	0	14.75	5.04	5.17	2.45	0	14.75
Lifeline100	7.28	7.35	3.08	.55	15.32	7.24	7.35	3.09	.55	15.32
Linkup	12.47	12.50	7.53	0	22.95	12.52	12.50	7.56	0	22.95
Autoenroll	.11	.00	.31	0	1	.10	.00	.30	0	1
Autoenroll2	.21	.00	.41	0	1	.21	.00	.41	0	1
Access Charge	.15	.14	.10	.03	.47	.15	.14	.10	.03	.47
White	.72	.80	.25	0	1	.71	.78	.25	0	1
Black	.16	.04	.23	0	.98	.17	.05	.24	0	.98
Native	.01	.00	.05	0	.99	.01	.00	.05	0	.99
Asian	.02	.00	.05	0	.88	.02	.00	.05	0	.73
Other	.09	.04	.11	0	.81	.09	.05	.11	0	.75
LCA	228480	64855	420043	203	3067685	239820	70989	431684	259	3067685
Median Income	43668	39442	17245	7371	175762	43312	38912	17154	12869	160223
Rural	.40	.22	.41	0	1	.36	.18	.39	0	1
MSA	.63	1.00	.48	0	1	.64	1.00	.48	0	1
State Rural	.25	.23	.14	0	.60	.25	.23	.14	0	.60
Competition	.16	.00	.36	0	1	.15	.00	.36	0	1
Elect PUC	.17	.00	.37	0	1	.17	.00	.38	0	1
Democrat PUC	34.74	33.33	27.12	0	100	34.78	33.33	27.34	0	100
Monthly50	13.59	13.50	2.50	8.55	21.75	13.57	13.50	2.51	8.55	21.75
Monthly100	15.82	15.67	2.75	10.25	23.95	15.79	15.67	2.75	10.25	23.95
Subsidy50	8.52	9.00	2.15	5.25	13.65	8.53	9.00	2.14	5.25	13.65
Subsidy100	12.50	9.00	63.12	3.55	1020.84	12.52	9.00	63.18	3.55	1020.84

*The Drop 50 sample excludes 521 wire centers which have 50 or fewer poor households.

Table 2 – First Approach

	ALL EXOGENOUS (1)		*Subsidy50* ENDOGENOUS (2)		*Lifeline50* ENDOGENOUS (3)		*Subsidy50, Linkup* ENDOGENOUS (4)		*Subsidy50, Autoenroll* ENDOGENOUS (5)	
Estimated Coefficients										
Lifeline50	-0.00198	*	-0.00303	***	-0.00383	***	-0.00509	**	-0.00273	**
	(0.00108)		(0.00113)		(0.00137)		(0.00257)		(0.00136)	
Linkup	-0.00049		-0.00068	*	-0.00070	*	0.00174		-0.00058	
	(0.00049)		(0.00039)		(0.00038)		(0.00272)		(0.00046)	
Autoenroll	0.02207	***	0.02476	***	0.02520	***	0.02508	***	0.01591	
	(0.00405)		(0.00311)		(0.00292)		(0.00858)		(0.02136)	
Black	-0.08562	***	-0.08251	***	-0.08643	***	-0.09480	***	-0.08415	***
	(0.01053)		(0.00878)		(0.00917)		(0.01661)		(0.00933)	
Native	-0.15741	***	-0.15233	***	-0.14934	***	-0.16978	**	-0.15290	***
	(0.05586)		(0.05097)		(0.05120)		(0.07031)		(0.05238)	
Asian	0.01806		0.00958		0.00972		0.01912		0.01093	
	(0.01865)		(0.01773)		(0.01830)		(0.02800)		(0.01910)	
Other	-0.08772	***	-0.08579	***	-0.09462	***	-0.08380	**	-0.08564	***
	(0.03250)		(0.02842)		(0.02991)		(0.03611)		(0.02912)	
ln(LCA)	0.00890	***	0.00895	***	0.00880	***	0.00879	***	0.00901	***
	(0.00123)		(0.00127)		(0.00128)		(0.00178)		(0.00124)	
Median Income	0.00094	***	0.00091	***	0.00087	***	0.00089	***	0.00093	***
	(0.00013)		(0.00013)		(0.00013)		(0.00020)		(0.00015)	
Rural	-0.02813	***	-0.02786	***	-0.02724	***	-0.02985	***	-0.02712	***
	(0.00502)		(0.00428)		(0.00433)		(0.00734)		(0.00454)	
MSA	0.01141	***	0.01142	***	0.01089	***	0.01308	***	0.01204	***
	(0.00412)		(0.00409)		(0.00406)		(0.00407)		(0.00418)	
constant	-0.18396	***	-0.17549	***	-0.16623	***	-0.19332	***	-0.18008	***
	(0.02082)		(0.02060)		-0.0204		(0.03988)		-0.02373	
Functions of Estimated Coefficients										
Elasticity-Lifeline	-0.01025	*	-0.01564	***	-0.01980	***	-0.02634	**	-0.01411	**
	(0.00562)		(0.00583)		(0.00707)		(0.01331)		(0.00702)	
Elasticity-Linkup	-0.00611		-0.00848	*	-0.00872	*	0.02172		-0.00720	
	(0.00618)		(0.00482)		(0.00472)		(0.03399)		(0.00572)	
Discount	0.24729		0.22427		0.18224		-0.34108		0.21104	
	(0.29782)		(0.15809)		(0.12408)		(0.48871)		(0.17394)	
Diagnostic Statistics										
R2	0.48		0.45		0.45		0.39		0.45	
J-statistic (d.f.)			1.36 (3)		1.26 (3)		1.12 (2)		1.12 (2)	
F-stat Lifeline50 (d.f.)			14.16 (4)	***	16.80 (4)	***	14.05 (4)	***	15.27 (4)	
F-stat Linkup (d.f.)							0.33 (4)			
F-stat Auto (d.f.)									1.44 (4)	
Hausman (ALL) vs. (1)			20.05 (12)	*	22.86 (12)	**	6.21 (12)		24.78 (12)	**
Hausman (ALL) vs. (2)					5.60 (12)		3.98 (12)		13.94 (12)	

Table 2 (cont.)

	Lifeline50, Linkup ENDOGENOUS (6)		*Lifeline50, Autoenroll* ENDOGENOUS (7)		*Subsidy50 Linkup,Autoenroll* ENDOGENOUS (8)		*ALL* ENDOGENOUS (9)	
Estimated Coefficients								
Lifeline50	-0.00648	*	-0.00388	***	-0.00391		-0.00128	
	(0.00339)		(0.00142)		(0.00654)		(0.01016)	
Linkup	0.00208		-0.00049		0.00066		-0.00348	
	(0.00296)		(0.00048)		(0.00658)		(0.01050)	
Autoenroll	0.02553	**	0.00540		0.01919		-0.00901	
	(0.01096)		(0.03001)		(0.04592)		(0.08118)	
Black	-0.10160	***	-0.09326	***	-0.09009	***	-0.08255	**
	(0.01943)		(0.01231)		(0.02677)		(0.03856)	
Native	-0.17287	**	-0.15515	***	-0.16319	**	-0.13349	*
	(0.07450)		(0.05375)		(0.07437)		(0.07509)	
Asian	0.02209		0.01533		0.01515		0.00412	
	(0.03096)		(0.02235)		(0.03265)		(0.04440)	
Other	-0.09873	**	-0.10179	***	-0.08330	***	-0.09877	***
	(0.03897)		(0.03121)		(0.03212)		(0.03386)	
ln(LCA)	0.00844	***	0.00871	***	0.00887	***	0.00892	***
	(0.00197)		(0.00128)		(0.00164)		(0.00231)	
Median Income	0.00087	***	0.00091	***	0.00091	***	0.00092	***
	(0.00020)		(0.00015)		(0.00018)		(0.00017)	
Rural	-0.03039	***	-0.02574	***	-0.02822	**	-0.02065	
	(0.00800)		(0.00486)		(0.01260)		(0.02034)	
MSA	0.01207	***	0.01167	***	0.01310	***	0.01196	*
	(0.00454)		(0.00402)		(0.00385)		(0.00671)	
constant	-0.18166	***	-0.16673	***	-0.18780	***	-0.14766	**
	(0.03433)		(0.01959)		(0.04265)		(0.07095)	
Functions of Estimated Coefficients								
Elasticity-Lifeline	-0.03350	*	-0.02008	***	-0.02019		-0.00660	
	(0.01751)		(0.00737)		(0.03383)		(0.05251)	
Elasticity-Linkup	0.02606		-0.00612		0.00820		-0.04356	
	(0.03696)		(0.00595)		(0.08221)		(0.13124)	
Discount	-0.32174		0.12615		-0.16796		2.72995	
	(0.37790)		(0.13957)		(1.41917)		(29.76417)	
Diagnostic Statistics								
R2	0.36		0.44		0.43		0.34	
J-statistic (d.f.)	0.41 (2)		0.45 (2)		1.35 (1)		0.12 (1)	
F-stat Lifeline50 (d.f.)	13.75 (4)	***	13.96 (4)	***	14.60 (4)	***	12.16 (4)	***
F-stat Linkup (d.f.)	0.80 (4)				0.32 (4)		0.68 (4)	
F-stat Auto (d.f.)			1.08 (4)		1.44 (4)		1.10 (4)	
Hausman (ALL) vs. (1)	5.13 (12)		16.46 (12)		12.68 (12)		3.56 (12)	
Hausman (ALL) vs. (2)	11.21 (12)		14.35 (12)		3.09 (12)		9.29 (12)	

Table 3.a - Robustness to Price Specification Changes (*Subsidy50* ENDOGENOUS)

	Lifeline100 (1)		*Autoenroll2* (2)		NO *Autoenroll* (3)		*Access* (4)		LOGARITHM (5)	
Estimated Coefficients										
Price									-0.02298	***
									(0.00648)	
Lifeline	-0.00173	*	-0.00200	**	-0.00231	**	-0.00277	**	5.86014	
	(0.00102)		(0.00094)		(0.00115)		(0.00118)		(3.91241)	
Linkup	-0.00079	*	-0.00051		-0.00046		-0.00065	*		
	(0.00041)		(0.00040)		(0.00042)		(0.00039)			
Autoenroll	0.02788	***	0.02685	***			0.02431	***	0.02684	***
	(0.00419)		(0.00440)				(0.00342)		(0.00289)	
Access							0.02245			
							(0.04351)			
Black	-0.07708	***	-0.07980	***	-0.08624	***	-0.08276	***	-0.08290	***
	(0.00810)		(0.00929)		(0.00969)		(0.00875)		(0.00867)	
Native	-0.14688	***	-0.16061	***	-0.15373	***	-0.15589	***	-0.15540	***
	(0.05370)		(0.04858)		(0.05311)		(0.05052)		(0.05161)	
Asian	0.01524		-0.01824		0.01701		0.01628		0.01047	
	(0.01703)		(0.01857)		(0.01888)		(0.02064)		(0.01766)	
Other	-0.08310	***	-0.10587	***	-0.08480	***	-0.08916	***	-0.08346	***
	(0.03196)		(0.02115)		(0.02939)		(0.03258)		(0.02731)	
ln(LCA)	0.00914	***	0.00921	***	0.00927	***	0.00893	***	0.00884	***
	(0.00121)		(0.00121)		(0.00120)		(0.00126)		(0.00121)	
Median Income	0.00097	***	0.00091	***	0.00098	***	0.00093	***	0.00091	***
	(0.00013)		(0.00013)		(0.00013)		(0.00013)		(0.00012)	
Rural	-0.02763	***	-0.02893	***	-0.02663	***	-0.02776	***	-0.02808	***
	(0.00432)		(0.00405)		(0.00491)		(0.00421)		(0.00420)	
MSA	0.01274	***	0.01087	***	0.01199	***	0.01086	***	0.01124	***
	(0.00406)		(0.00389)		(0.00410)		(0.00385)		(0.00390)	
constant	-0.18515	***	-0.18616	***	-0.18803	***	-0.18067	***	-0.11409	***
	(0.02134)		(0.02012)		(0.02047)		(0.02293)		(0.03301)	
Functions of Estimated Coefficients										
Elasticity-Lifeline	-0.01274	*	-0.01033	**	-0.01193	**	-0.01430	**	-0.01590	***
	(0.00749)		(0.00485)		(0.00596)		(0.00608)		(0.00553)	
Elasticity-Linkup	-0.00983	*	-0.00636		-0.00578		-0.00817	*	-0.00656	*
	(0.00511)		(0.00503)		(0.00530)		(0.00489)		(0.00356)	
Discount	0.45351		0.25454		0.20036		0.23640		0.17064	
	(0.32123)		(0.22529)		(0.20354)		(0.17255)		(0.11393)	
Diagnostic Statistics										
R2	0.45		0.46		0.44		0.45		0.45	
J-statistic (d.f.)	3.27 (3)		1.55 (3)		1.28 (3)		1.25 (3)		1.37 (3)	
F-stat Lifeline(d.f.)	14.50 (4)	***	15.21 (4)	***	15.27 (4)	***	13.73 (4)	***	14.16 (4)	***
Hausman (ALL) vs. (1)	25.24 (12)	**	12.77 (12)		27.92 (11)	***	25.43 (13)	**	16.45 (12)	
N	6570		6596		6596		6596		6596	

* 90% confidence   ** 95% confidence   *** 99% confidence

Table 4.a - Robustness to Sample Changes (*Subsidy50* ENDOGENOUS)

	DROP CALIFORNIA (1)		DROP *N* <=100 (2)		ALL LOCATION (3)		DROP 100% LOCATION (4)	
Estimated Coefficients								
Lifeline50	-0.00239	**	-0.00263	**	-0.00324	***	-0.00277	***
	(0.00109)		(0.00104)		(0.00123)		(0.00107)	
Linkup	-0.00056		-0.00071	*	-0.00060		-0.00068	*
	(0.00040)		(0.00038)		(0.00039)		(0.00039)	
Autoenroll	0.02491	***	0.02104	***	0.02621	***	0.02212	***
	(0.00303)		(0.00369)		(0.00332)		(0.00286)	
Black	-0.08219	***	-0.07869	***	-0.08838	***	-0.07700	***
	(0.00938)		(0.00846)		(0.00921)		(0.00864)	
Native	-0.15415	***	-0.13957	**	-0.16157	***	-0.14267	***
	(0.05050)		(0.05599)		(0.04630)		(0.05080)	
Asian	0.01887		0.00938		0.01250		0.02458	
	(0.02519)		(0.01660)		(0.01831)		(0.02111)	
Other	-0.12287	***	-0.08778	***	-0.08767	***	-0.08289	***
	(0.02191)		(0.02886)		(0.02923)		(0.02727)	
ln(LCA)	0.00948	***	0.00902	***	0.00890	***	0.00907	***
	(0.00124)		(0.00130)		(0.00120)		(0.00124)	
Median Income	0.00096	***	0.00093	***	0.00089	***	0.00088	***
	(0.00014)		(0.00013)		(0.00014)		(0.00014)	
Rural	-0.02868	***	-0.03525	***	-0.02184	***	-0.03058	***
	(0.00431)		(0.00453)		(0.00377)		(0.00423)	
MSA	0.00877	**	0.01052	***	0.01160	***	0.00986	**
	(0.00391)		(0.00385)		(0.00438)		(0.00395)	
constant	-0.18503	***	-0.17764	***	-0.17431	***	-0.17932	***
	(0.02042)		(0.02119)		(0.01937)		(0.02073)	
Functions of Estimated Coefficients								
Elasticity-Lifeline	-0.01315	**	-0.01359	**	-0.01680	***	-0.01441	***
	(0.00599)		(0.00539)		(0.00637)		(0.00559)	
Elasticity-Linkup	-0.00791		-0.00883	*	-0.00752		-0.00849	*
	(0.00559)		(0.00475)		(0.00492)		(0.00482)	
Discount	0.23404		0.26857		0.18573		0.24552	
	(0.19654)		(0.18069)		(0.14928)		(0.16927)	
Diagnostic Statistics								
R2	0.45		0.48		0.43		0.43	
J-statistic (d.f.)	2.07 (3)		1.32 (3)		1.54 (3)		1.15 (3)	
F-stat Lifeline(d.f.)	12.95 (4)	***	14.78 (4)	***	13.92 (4)	***	13.81 (4)	***
Hausman (ALL) vs. (1)	18.91 (12)	*	17.20 (12)		6.32 (12)		22.96 (12)	**
N	6138		5884		7117		6012	

Table 5 – Second Approach

	Full Sample (7117 obs.)			Drop 50 Sample (6596 obs.)		
	ALL	*Subsidy50*	*Lifeline50*	ALL	*Subsidy50*	*Lifeline50*
	EXOGENOUS	ENDOGENOUS	ENDOGENOUS	EXOGENOUS	ENDOGENOUS	ENDOGENOUS
	(1)	(2)	(3)	(4)	(5)	(6)
Estimated Coefficients						
Lifeline50	-0.0021	-0.0037 ***	-0.0043 ***	-0.0023 *	-0.0036 ***	-0.0043 ***
	(0.0014)	(0.0013)	(0.0015)	(0.0013)	(0.0012)	(0.0014)
Linkup	-0.0005	-0.0008	-0.0008 *	-0.0006	-0.0008 *	-0.0009 *
	(0.0006)	(0.0005)	(0.0005)	(0.0006)	(0.0005)	(0.0005)
Autoenroll	0.0342 ***	0.0332 ***	0.0330 ***	0.0354 ***	0.0342 ***	0.0341 ***
	(0.0076)	(0.0068)	(0.0072)	(0.0076)	(0.0063)	(0.0066)
Black	-0.0912 ***	-0.0863 ***	-0.0894 ***	-0.0876 ***	-0.0821 ***	-0.0855 ***
	(0.0116)	(0.0092)	(0.0090)	(0.0123)	(0.0087)	(0.0087)
Native	-0.1677 ***	-0.1663 ***	-0.1670 **	-0.1634 ***	-0.1576 **	-0.1575 **
	(0.0515)	(0.0603)	(0.0725)	(0.0568)	(0.0618)	(0.0723)
Asian	0.5539	0.5267	0.5135	0.5860	0.5106	0.5027
	(1.24E+14)	(3.32E+13)	(6.48E+12)	(1.82E+15)	(7.16E+11)	(2.95E+11)
Other	-0.0877 **	-0.0731 **	-0.0779 ***	-0.0865 **	-0.0742 **	-0.0790 ***
	(0.0388)	(0.0315)	(0.0295)	(0.0386)	(0.0305)	(0.0288)
ln(LCA)	0.0089 ***	0.0087 ***	0.0086 ***	0.0088 ***	0.0087 ***	0.0085 ***
	(0.0021)	(0.0026)	(0.0024)	(0.0022)	(0.0023)	(0.0023)
Median Income	0.0017 ***	0.0017 ***	0.0016 ***	0.0017 ***	0.0017 ***	0.0016 ***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Rural	-0.0226 ***	-0.0231 ***	-0.0229 ***	-0.0267 ***	-0.0274 ***	-0.0273 ***
	(0.0056)	(0.0066)	(0.0064)	(0.0063)	(0.0070)	(0.0067)
MSA	0.0080	0.0091 *	0.0089 *	0.0087	0.0089 *	0.0088 *
	(0.0055)	(0.0052)	(0.0048)	(0.0055)	(0.0050)	(0.0047)
constant	-0.2049 ***	-0.1906 ***	-0.1828 ***	-0.2027 ***	-0.1902 ***	-0.1811 ***
	(0.0297)	(0.0331)	(0.0309)	(0.0309)	(0.0314)	(0.0301)
sigma_s	0.0552 ***	0.0528 ***	0.0532 ***	0.0563 ***	0.0539 ***	0.0544 ***
	(0.0101)	(0.0129)	(0.0126)	(0.0119)	(0.0109)	(0.0106)
correlation		0.1016	0.1533 **		0.0920	0.1443
		(0.0821)	(0.0756)		(0.0741)	(0.0987)
sigma_p		1.6675 ***	1.7481 ***		1.6516 ***	1.7435 ***
		(0.1501)	(0.1967)		(0.1455)	(0.2007)
Functions of Estimated Coefficients						
ElasticityLifeline	-0.0111	-0.0192 ***	-0.0221 ***	-0.0118 *	-0.0188 ***	-0.0220 ***
	(0.0071)	(0.0067)	(0.0077)	(0.0069)	(0.0062)	(0.0071)
ElasticityLinkup	-0.0066	-0.0095	-0.0099 *	-0.0075	-0.0105 *	-0.0110 *
	(0.0074)	(0.0061)	(0.0059)	(0.0073)	(0.0061)	(0.0059)
Discount	0.2453 **	0.2055 ***	0.1860 ***	0.2638 ***	0.2318 ***	0.2071 ***
	(0.1098)	(0.0254)	(0.0183)	(0.0976)	(0.0263)	(0.0185)
Diagnostic Statistics						
R2	0.47	0.47	0.47	0.49	0.49	0.49
J-statistic (d.f.)		3.75(3)	3.72(3)		2.46(3)	3.20(3)
F-stat Lifeline50 (d.f.)		12.95(4) ***	21.95(4) ***		13.35(4) ***	20.71(4) ***
Hausman(ALL) vs. All Exog.		5.86(11)	7.63(11)		5.56(11)	7.53(11)
Hausman(ALL) vs.Subs. Endog.			8.69(11)			10.51(11)

Table 6.a – Impact of Lifeline and Linkup on Low-Income Penetration

Policy Experiments based on *Subsidy50* ENDOGENOUS Estimates*

FIRST APPROACH, DROP 50 SAMPLE

Baseline autoenrollment policy	Actual	Zero in all states	Actual
Autoenrollment after policy change	Actual	Zero in all states	Zero in all states
Policy change		*Resulting decrease in penetration*	
Eliminate Lifeline and Linkup		3.8%	4.1%
		[1.2%, 6.5%]	[1.6%, 6.8%]
Eliminate Lifeline	2.3%	2.3%	
	[0.5%, 4.2%]	[0.5%, 4.2%]	
Eliminate Linkup	1.6%	1.6%	
	[-0.1%, 3.4%]	[-0.1%, 3.3%]	

Actual penetration in sample = 92.2% [95% Confidence Interval]

SECOND APPROACH, FULL SAMPLE

Baseline autoenrollment policy	Actual	Zero in all states	Actual
Autoenrollment after policy change	Actual	Zero in all states	Zero in all states
Baseline Expected Penetration	92.6%	92.3%	92.6%
Policy change		*Resulting decrease in penetration*	
Eliminate Lifeline and Linkup		3.9%	4.2%
		[1.6%, 6.2%]	[1.9%, 6.5%]
Eliminate Lifeline	2.4%	2.4%	
	[0.6%, 3.9%]	[0.6%, 4.0%]	
Eliminate Linkup	1.4%	1.4%	
	[-0.3%, 3.3%]	[-0.3%, 3.3%]	

Actual penetration in sample = 92.2% [95% Confidence Interval]

* Table 2, column 2; Table 5, column 2

Table 6.b - Impacts of Lifeline and Linkup on Low-Income Penetration

Policy Experiments based on *Lifeline50* ENDOGENOUS estimates**

FIRST APPROACH, DROP 50 SAMPLE

Baseline autoenrollment policy	Actual	Zero in all states	Actual
Autoenrollment after policy change	Actual	Zero in all states	Zero in all states
Policy change	*Resulting decrease in penetration*		
Eliminate Lifeline and Linkup		4.5%	4.8%
		[1.8%, 7.0%]	[2.2%, 7.3%]
Eliminate Lifeline	2.9%	2.9%	
	[1.0%, 5.4%]	[1.0%, 5.4%]	
Eliminate Linkup	1.6%	1.6%	
	[-0.5%, 3.0%]	[-0.5%, 3.0%]	

Actual penetration in sample = 92.2% [95% Confidence Interval]

SECOND APPROACH, FULL SAMPLE

Baseline autoenrollment policy	Actual	Zero in all states	Actual
Autoenrollment after policy change	Actual	Zero in all states	Zero in all states
Baseline Expected Penetration	92.6%	92.3%	92.6%
Policy change	*Resulting decrease in penetration*		
Eliminate Lifeline and Linkup		4.4%	4.7%
		[1.7%, 6.9%]	[2.0%, 7.2%]
Eliminate Lifeline	2.7%	2.8%	
	[0.8%, 4.8%]	[0.9%, 4.8%]	
Eliminate Linkup	1.5%	1.5%	
	[-0.3%, 3.4%]	[-0.4%, 3.4%]	

Actual penetration in sample = 92.2% [95% Confidence Interval]

** Table 2, column 3; Table 5, column 3

Appendix A – Moments for GMM estimation (second approach)

Let $N_{gl} \equiv X_g N_l$ denote the number of households in group g at location. Then

$$S_{gl} N_{gl} \sim Binomial(N_{gl}, \overline{S}_{gl})$$

with

$$\overline{S}_{gl} = \min\left\{1, \mathrm{e}^{\delta_l + \theta_g}\right\}$$

where δ_l is a location fixed effect and θ_g is a group fixed effect. Therefore, the first and second moments of the penetration rate at location l are

$$E\left\{S_{gl} \mid \delta_l\right\} = \overline{S}_{gl}$$

$$E\left\{S_{gl}{}^2 \mid \delta_l\right\} = \tfrac{1}{N_{gl}} \overline{S}_{gl} + \tfrac{N_{gl}-1}{N_{gl}} \overline{S}_{gl}^2$$

If

$$\delta_l = \psi_l + w_l$$

$$w_l \sim N(0, \sigma_w^2)$$

then

$$E\left\{S_{gl} \mid \psi_l\right\} = \int_{-\infty}^{\infty} \min\left\{1, e^{\psi_l + \theta_g + w_l}\right\} \tfrac{1}{\sigma_w} \phi(\tfrac{w_l}{\sigma_w}) dw_l$$

$$= e^{\psi_l + \theta_g} \int_{-\infty}^{-\psi_l - \theta_g} e^{w_l} \tfrac{1}{\sigma_w} \phi(\tfrac{w_l}{\sigma_w}) dw_l + \int_{-\psi_l - \theta_g}^{\infty} \tfrac{1}{\sigma_w} \phi(\tfrac{w_l}{\sigma_w}) dw_l$$

$$= e^{\psi_l + \theta_g + \frac{1}{2}\sigma_w^2} \Phi(\tfrac{-\psi_l - \theta_g}{\sigma_w} - \sigma_w) + 1 - \Phi(\tfrac{-\psi_l - \theta_g}{\sigma_w})$$

$$E\left\{S_{gl}{}^2 \mid \psi_l\right\} = \int_{-\infty}^{-\psi_l - \theta_g} \left[\tfrac{1}{N_{gl}} e^{\psi_l + \theta_g + w_l} + \tfrac{N_{gl}-1}{N_{gl}} e^{2(\psi_l + \theta_g + w_l)} \right] \tfrac{1}{\sigma_w} \phi(\tfrac{w_l}{\sigma_w}) dw_l + \int_{-\psi_l - \theta_g}^{\infty} \tfrac{1}{\sigma_w} \phi(\tfrac{w_l}{\sigma_w}) dw_l$$

$$= \tfrac{1}{N_{gl}} e^{\psi_l + \theta_g + \frac{1}{2}\sigma_w^2} \Phi(\tfrac{-\psi_l - \theta_g}{\sigma_w} - \sigma_w) + \tfrac{N_{gl}-1}{N_{gl}} e^{2(\psi_l + \theta_g + \sigma_w^2)} \Phi(\tfrac{-\psi_l - \theta_g}{\sigma_w} - 2\sigma_w) + 1 - \Phi(\tfrac{-\psi_l - \theta_g}{\sigma_w})$$

where $\phi(\cdot)$ is the standard normal density, and $\Phi(\cdot)$ the standard normal cumulative distribution function.

Furthermore, if $\theta_{g'} > \theta_g$, then

$$E\{S_{gl} S_{g'l} \mid \psi_l\} = e^{2\psi_l + \theta_g + \theta_{g'}} \int_{-\infty}^{-\psi_l - \theta_{g'}} e^{2w_l} \tfrac{1}{\sigma_w} \phi(\tfrac{w_l}{\sigma_w}) dw_l + e^{\psi_l + \theta_g} \int_{-\psi_l - \theta_{g'}}^{-\psi_l - \theta_g} e^{w_l} \tfrac{1}{\sigma_w} \phi(\tfrac{w_l}{\sigma_w}) dw_l + \int_{-\psi_l - \theta_g}^{\infty} \tfrac{1}{\sigma_w} \phi(\tfrac{w_l}{\sigma_w}) dw_l$$

$$= e^{2\psi_l + \theta_g + \theta_{g'} + 2\sigma_w^2} \Phi(\tfrac{-\psi_l - \theta_{g'}}{\sigma_w} - 2\sigma_w) + e^{\psi_l + \theta_g + \frac{\sigma_w^2}{2}} [\Phi(\tfrac{-\psi_l - \theta_g}{\sigma_w} - \sigma_w) - \Phi(\tfrac{-\psi_l - \theta_{g'}}{\sigma_w} - \sigma_w)] + 1 - \Phi(\tfrac{-\psi_l - \theta_g}{\sigma_w})$$

Aggregating over groups

$$E\left\{S_l \mid \psi_l\right\} = \sum_g X_{gl} E\left\{S_{gl} \mid \psi_l\right\}$$

$$E\left\{S_l^2 \mid \psi_l\right\} = \sum_g X_{gl}^2 E\left\{S_{gl}^2 \mid \psi_l\right\} + 2 \sum_g \sum_{g'>g} X_{gl} X_{g'l} E\{S_{gl} S_{g'l} \mid \psi_l\}$$

Thus these formulas enable calculation of the first two moments conditional on observable characteristics by integrating over a binomial sampling error and a normally distributed error for each location associated with a reduced form penetration equation. In particular, if the structural model is

$$\delta_l = \boldsymbol{\theta}_x \cdot \mathbf{x}_l + \theta_y y_l + u_l$$

$$y_l = \boldsymbol{\theta}_z \cdot \mathbf{z}_l + v_l$$

$$\left(u_l, v_l\right) \sim N(0, \begin{bmatrix} \sigma_u^2 & \sigma_{uv} \\ \sigma_{uv} & \sigma_v^2 \end{bmatrix})$$

$$\sigma_{uv} \equiv \rho \sigma_u \sigma_v$$

then

$$\psi_l = \boldsymbol{\theta}_x \cdot \mathbf{x}_l + \theta_y \boldsymbol{\theta}_z \cdot \mathbf{z}_l$$

$$w_l = u_l + \theta_y v_l$$

$$\sigma_w^2 \equiv \sigma_u^2 + \theta_y^2 \sigma_v^2 + 2\theta_y \sigma_{uv}$$

and it is straightforward to calculate expectations for S_l and S_l^2 conditional on $\left(\mathbf{x}_l, \mathbf{z}_l\right)$.

It is straightforward to calculate moments for y_l.

An additional moment of interest is $E\{y_l S_l \mid x_l, \mathbf{z}_l\}$. This calculation uses

$$\left(w_l, v_l\right) \sim N(0, \begin{bmatrix} \sigma_w^2 & \theta_y \sigma_v^2 + \sigma_{uv} \\ \theta_y \sigma_v^2 + \sigma_{uv} & \sigma_v^2 \end{bmatrix})$$

Let $f(v_l \mid w_l)$ denote the conditional density. Then

$$
\begin{aligned}
E\{y_l S_{gl} \mid \psi_l, \mathbf{z}_l\} &= \tfrac{1}{\sigma_w} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} (\boldsymbol{\theta}_z \cdot \mathbf{z}_l + v_l) f(v_l \mid w_l) dv_l \right] \min\left\{1, e^{\psi_l + \theta_g + w_l}\right\} \phi(\tfrac{w_l}{\sigma_w}) dw_l \\
&= \int_{-\infty}^{-\psi_l - \theta_g} \left(\boldsymbol{\theta}_z \cdot \mathbf{z}_l + \tfrac{\theta_y \sigma_v^2 + \sigma_{uv}}{\sigma_w^2} w_l \right) e^{\psi_l + \theta_g + w_l} \tfrac{1}{\sigma_w} \phi(\tfrac{w_l}{\sigma_w}) dw_l + \boldsymbol{\theta}_z \cdot \mathbf{z}_l \left[1 - \Phi(\tfrac{-\psi_l - \theta_g}{\sigma_w}) \right] + \int_{-\psi_l - \theta_g}^{\infty} \tfrac{\theta_y \sigma_v^2 + \sigma_{uv}}{\sigma_w^3} w_l \phi(\tfrac{w_l}{\sigma_w}) dw_l \\
&= \int_{-\infty}^{-\psi_l - \theta_g} \left(\boldsymbol{\theta}_z \cdot \mathbf{z}_l + \tfrac{\theta_y \sigma_v^2 + \sigma_{uv}}{\sigma_w^2} w_l \right) e^{\psi_l + \theta_g + w_l} \tfrac{1}{\sigma_w} \phi(\tfrac{w_l}{\sigma_w}) dw_l + \boldsymbol{\theta}_z \cdot \mathbf{z}_l \left[1 - \Phi(\tfrac{-\psi_l - \theta_g}{\sigma_w}) \right] + \tfrac{\theta_y \sigma_v^2 + \sigma_{uv}}{\sigma_w \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{-\psi_l - \theta_g}{\sigma_w}\right)^2} \\
&= e^{\psi_l + \theta_g} \left[\left(\boldsymbol{\theta}_z \cdot \mathbf{z}_l \right) \int_{-\infty}^{-\psi_l - \theta_g} e^{w_l} \phi(\tfrac{w_l}{\sigma_w}) \tfrac{1}{\sigma_w} dw_l + \tfrac{\theta_y \sigma_v^2 + \sigma_{uv}}{\sigma_w^2} \int_{-\infty}^{-\psi_l - \theta_g} w_l e^{w_l} \phi(\tfrac{w_l}{\sigma_w}) \tfrac{1}{\sigma_w} dw_l \right] \\
&\quad + \boldsymbol{\theta}_z \cdot \mathbf{z}_l \left[1 - \Phi(\tfrac{-\psi_l - \theta_g}{\sigma_w}) \right] + \tfrac{\theta_y \sigma_v^2 + \sigma_{uv}}{\sigma_w \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{-\psi_l - \theta_g}{\sigma_w}\right)^2} \\
&= e^{\psi_l + \theta_g + \frac{\sigma_w^2}{2}} \left\{ \left(\boldsymbol{\theta}_z \cdot \mathbf{z}_l \right) \Phi(\tfrac{-\psi_l - \theta_g - \sigma_w^2}{\sigma_w}) + \left(\theta_y \sigma_v^2 + \sigma_{uv} \right) \left[\Phi(\tfrac{-\psi_l - \theta_g - \sigma_w^2}{\sigma_w}) - \tfrac{1}{\sigma_w \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{-\psi_l - \theta_g - \sigma_w^2}{\sigma_w}\right)^2} \right] \right\} \\
&\quad + \boldsymbol{\theta}_z \cdot \mathbf{z}_l \left[1 - \Phi(\tfrac{-\psi_l - \theta_g}{\sigma_w}) \right] + \tfrac{\theta_y \sigma_v^2 + \sigma_{uv}}{\sigma_w \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{-\psi_l - \theta_g}{\sigma_w}\right)^2}
\end{aligned}
$$

and aggregating over groups gives

$$
E\{y_l S_l \mid \psi_l, \mathbf{z}_l\} = \sum_g X_{gl} E\{y_l S_{gl} \mid \psi_l, \mathbf{z}_l\}
$$

This formula is used to construct the covariance moment conditional on $\left(\mathbf{x}_l, \mathbf{z}_l\right)$.

Appendix B – Estimating Lifeline and Linkup expense in our sample

Because the sample of 7,117 wire centers employed in the study does not cover the whole country, it is necessary to estimate the total cost of the Lifeline and Linkup programs for the areas included in the sample. Our data do not include all of the households in a state for two reasons: 1) RBOCs do not typically serve the entire state; and 2) we drop wire centers for which we could not identify a unique price.

We employ four main sources of data to estimate the cost of the Lifeline and Linkup programs for our sample; 1) FCC ARMIS database, containing data on the number of Lifeline lines in each study area (each company in each state); 2) FCC "Monitoring Report" information on the federal Lifeline and Link-Up subsidies to each study area in each state; 3) FCC (2003) estimates of the number of households eligible for the Lifeline and Linkup programs in each state in 2000; and 4) the census variable P92 (Poverty Status in 1999 of households by household type) to determine the number of poor households in each wire center and study area.

Because households above the poverty level are eligible to receive Lifeline and Linkup subsidies in several states (e.g., California households with incomes below 150% of the poverty line are eligible for Lifeline and Linkup subsidies), the actual number of Lifeline subscribers may overestimate the number of households receiving Lifeline and Linkup subsidies in our sample of households below the poverty level. To estimate the number of households below the poverty level receiving Lifeline and Linkup subsidies, we compared the FCC's estimate of the number of households eligible for the Lifeline subsidy (*Eligible HH*) with the actual number of Lifeline recipients (*Lifelines*) and the number of households below the poverty level (*Pov HH*). In study areas where households that were eligible for or receiving the Lifeline subsidies exceeded the number of poor households, we deflated the number of Lifeline lines with the following weight:

$$w = Pov\,HH \,/\, \left(max(\,Eligible\,HH\,, Lifelines\,) \right).$$

In cases where the number of households below the poverty line in a study area exceeds both the number of eligible households and the number of Lifeline lines, we assume that all households receiving the Lifeline subsidy had incomes below the poverty level (*w*=1). Weighted Lifeline lines equals the product of Lifeline lines in a study area and *w*. The same methodology is used to determine weighted Linkup dollars.

Federal and state per line subsidies for Lifeline are calculated for each state as follows:

$$Subsidy50 = Monthly50 - Lifeline50$$

$$Federal50 = Min\left[\left(\$1.75 + SLC + \frac{Monthly50 - Lifeline50 - (SLC + \$1.75)}{3} \right), \$7 \right]$$

$$State50 = Monthly50 - Lifeline50 - Federal50\,,$$

where *SLC* equals the federal subscriber line charge.[1] The Lifeline subsidy in each wire center (and the amount corresponding to the Federal and State governments) equals the product of the number Lifeline lines allocated to a wire center and the per-line subsidy. The total cost of the Lifeline program in our sample equals the sum of the subsidies in each wire center.

We allocate federal Linkup dollars to each wire center using the product of the share of state poor households corresponding to each wire center and the annual federal Linkup in the state. federal and state per line connection subsidies are calculated as follow:

$$SubsidyLU = Hookup - Linkup$$

$$FederalLU = Min(.50 * Hookup, 30)$$

$$StateLU = Linkup - FederalLU$$

We estimate number of Linkup households in our data as the ratio of Federal dollars allocated to our data (from above) to Federal per line subsidy. The Federal and state Linkup subsidies per wire center equal the product of the number of estimated Linkups in each wire center and the per-line subsidies. The estimated total cost of the Linkup program in our sample equals the sum of the estimated wire-center subsidies.

[1] With the exception of the District of Columbia, the federal residential SLC equaled $3.50 in all states on January 1, 2000. The SLC equaled $3.32 in the District of Columbia.

Appendix C – Additional Tables

Table 3.b - Robustness to Price Specification Changes (*Lifeline* ENDOGENOUS)

	Lifeline100 (1)		*Autoenroll2* (2)		NO *Autoenroll* (3)		*Access* (4)		LOGARITHM (5)	
Estimated Coefficients										
Price									-0.02706	***
									(0.00738)	
Lifeline	-0.00368	***	-0.00306	**	-0.00384	***	-0.00368	***	7.02253	
	(0.00141)		(0.00126)		(0.00146)		(0.00141)		(4.54815)	
Linkup	-0.00088	**	-0.00050		-0.00045		-0.00066	*		
	(0.00037)		(0.00039)		(0.00041)		(0.00038)			
Autoenroll	0.03296	***	0.02794	***			0.02489	***	0.02750	***
	(0.00414)		(0.00441)				(0.00316)		(0.00281)	
Access							0.02675			
							(0.04337)			
Black	-0.08330	***	-0.08385	***	-0.09424	***	-0.08701	***	-0.08646	***
	(0.00789)		(0.00997)		(0.01075)		(0.00908)		(0.00874)	
Native	-0.15928	***	-0.16094	***	-0.15732	***	-0.15131	***	-0.15321	***
	(0.05257)		(0.04788)		(0.05288)		(0.05072)		(0.05169)	
Asian	0.00991		-0.02750		0.01691		0.01730		0.00996	
	(0.01865)		(0.01837)		(0.02079)		(0.02084)		(0.01819)	
Other	-0.10309	***	-0.11056	***	-0.10264	***	-0.09907	***	-0.09209	***
	(0.03381)		(0.02016)		(0.03077)		(0.03486)		(0.02888)	
ln(LCA)	0.00909	***	0.00894	***	0.00872	***	0.00869	***	0.00871	***
	(0.00122)		(0.00120)		(0.00129)		(0.00125)		(0.00122)	
Median Income	0.00090	***	0.00084	***	0.00092	***	0.00089	***	0.00088	***
	(0.00013)		(0.00013)		(0.00012)		(0.00013)		(0.00012)	
Rural	-0.02710	***	-0.02819	***	-0.02572	***	-0.02755	***	-0.02756	***
	(0.00453)		(0.00407)		(0.00494)		(0.00425)		(0.00424)	
MSA	0.01247	***	0.01128	***	0.01170	***	0.01083	***	0.01080	***
	(0.00399)		(0.00389)		(0.00401)		(0.00387)		(0.00383)	
constant	-0.16313	***	-0.17431	***	-0.16754	***	-0.17063	***	-0.09105	**
	(0.02084)		(0.01976)		(0.01979)		(0.02270)		(0.03748)	
Functions of Estimated Coefficients										
Elasticity-Lifeline	-0.02705	***	-0.01584	**	-0.01987	***	-0.01903	***	-0.01972	***
	(0.01036)		(0.00650)		(0.00753)		(0.00729)		(0.00659)	
Elasticity-Linkup	-0.01096	**	-0.00626		-0.00566		-0.00825	*	-0.00679	*
	(0.00458)		(0.00490)		(0.00513)		(0.00473)		(0.00359)	
Discount	0.23824	*	0.16348		0.11792		0.17930		0.14240	
	(0.13579)		(0.14958)		(0.12523)		(0.12640)		(0.09222)	
Diagnostic Statistics										
R2	0.44		0.46		0.44		0.45		0.45	
J-statistic (d.f.)	1.06 (3)		1.10 (3)		0.47 (3)		1.23 (3)		1.02 (3)	
F-stat Lifeline(d.f.)	12.64 (4)	***	16.14 (4)	***	13.96 (4)	***	16.49 (4)	***	16.80 (4)	***
Hausman (ALL) vs. (1)	15.25 (12)		18.37 (12)		18.02 (11)	*	27.16 (13)	**	28.26 (12)	***
N	6570		6596		6596		6596		6596	

* 90% confidence ** 95% confidence *** 99% confidence

Table 4.b - Robustness to Sample Changes (*Lifeline50* ENDOGENOUS)

	DROP CALIFORNIA (1)		DROP *N* <=100 (2)		ALL LOCATION (3)		DROP 100% LOCATION (4)	
Estimated Coefficients								
Lifeline50	-0.00333	**	-0.00351	***	-0.00396	***	-0.00351	**
	(0.00138)		(0.00129)		(0.00142)		(0.00136)	
Linkup	-0.00061		-0.00072	*	-0.00062		-0.00069	*
	(0.00039)		(0.00037)		(0.00039)		(0.00038)	
Autoenroll	0.02531	***	0.02179	***	0.02641	***	0.02276	***
	(0.00288)		(0.00325)		(0.00312)		(0.00283)	
Black	-0.08620	***	-0.08301	***	-0.09170	***	-0.08033	***
	(0.00987)		(0.00882)		(0.00942)		(0.00907)	
Native	-0.14979	***	-0.13748	**	-0.16018	***	-0.13927	***
	(0.05102)		(0.05606)		(0.04623)		(0.05105)	
Asian	0.01892		0.00996		0.01148		0.02541	
	(0.02469)		(0.01718)		(0.01877)		(0.02184)	
Other	-0.13155	***	-0.09739	***	-0.09527	***	-0.09027	***
	(0.02192)		(0.03025)		(0.03038)		(0.02884)	
ln(LCA)	0.00924	***	0.00880	***	0.00883	***	0.00897	***
	(0.00125)		(0.00131)		(0.00121)		(0.00124)	
Median Income	0.00091	***	0.00090	***	0.00086	***	0.00085	***
	(0.00014)		(0.00013)		(0.00014)		(0.00014)	
Rural	-0.02777	***	-0.03416	***	-0.02144	***	-0.03003	***
	(0.00437)		(0.00457)		(0.00382)		(0.00426)	
MSA	0.00860	**	0.00999	***	0.01114	***	0.00915	**
	(0.00389)		(0.00387)		(0.00424)		(0.00392)	
constant	-0.17316	***	-0.16715	***	-0.16662	***	-0.17147	***
	(0.02039)		(0.02097)		(0.01932)		(0.02056)	
Functions of Estimated Coefficients								
Elasticity-Lifeline	-0.01834	**	-0.01813	***	-0.02055	***	-0.01828	**
	(0.00762)		(0.00665)		(0.00738)		(0.00711)	
Elasticity-Linkup	-0.00861		-0.00902	*	-0.00773		-0.00866	*
	(0.00546)		(0.00461)		(0.00487)		(0.00471)	
Discount	0.18281		0.20572		0.15616		0.19747	
	(0.14227)		(0.13376)		(0.12023)		(0.13477)	
Diagnostic Statistics								
R2	0.45		0.48		0.43		0.42	
J-statistic (d.f.)	1.56 (3)		1.44 (3)		1.24 (3)		1.24 (3)	
F-stat Lifeline(d.f.)	14.25 (4)	***	16.95 (4)	***	17.25 (4)	***	14.97 (4)	***
Hausman (ALL) vs. (1)	20.35 (12)	*	21.99 (12)	**	7.43 (12)		20.97 (12)	*
N	6138		5884		7117		6012	