

Constrained Monopoly Pricing with Endogenous Participation^{*}

Gabriel Basaluzzo[†]

Eugenio J. Miravete[‡]

June 23, 2005

Abstract

We present a flexible model of monopoly nonlinear pricing with endogenous participation decisions of heterogeneous consumers. We make use of the moments that define the few self-selecting tariff options that are commonly used to implement the optimal nonlinear tariff to estimate how demand and cost variables affect the pricing strategies offered by incumbent monopolists in several early U.S. local cellular telephone markets through the different elements of the theoretical model: marginal costs, average price sensitivity of demand, indexing parameters governing the distribution of the two-dimensional type components, support of the distribution of types, and costs associated to the commercialization of tariff options. The sources of identification are the position and shape of each tariff offered by monopolists, the actual number and features of the tariff options used to implement them, as well as a measure of market penetration in each cellular market during the first and last quarter of monopoly regime. We use our model and the structural estimates to provide a performance comparison (profit+welfare) of *nonlinear* tariffs relative to *linear* (uniform), optimal *two-part*, Coasian *marginal cost-plus fixed fee*, and *flat* tariffs. We furthermore evaluate the potential welfare gains of implementing universal service requirements.

Keywords: Nonlinear Pricing; Endogenous Participation; Quantity Underprovision; Nonlinear *vs.* Restricted Pricing; Universal Service.

JEL Codes: C63, D43, D82, L96

^{*} We thank Toker Doganoglu, Gautam Gowrisankaran, Phil Haile, Bo Honoré, Elena Krasnokutskaya, Thierry Magnac, Charles Manski, Matthew Shum, and Kenneth Wolpin for their comments and suggestions. We also thank participants at the 2004 Kiel-Munich Workshop on the Economics of Information and Network Industries; the 2004 IO Day conference at the NYU-Stern School of Business; the 2004 conference on “The Economics of Electronic Communication Markets at IDEI-Toulouse; the 2005 Workshop on Mathematical Structures in Economics and Econometrics: “Aggregation and Disaggregation: Characterization and Identification of Collective Demand,” at the Banff International Research Station for Mathematical Innovation and Discovery; the 2005 Society of Computational Economics Conference on Computing in Economics and Finance at George Washington University; as well as seminar audiences at Universidad Carlos III, ITAM, Northwestern University, University of Pennsylvania, Purdue University, University of Southern California, Texas A&M University, and University of Toronto. For their logistic support, we wish to thank the Experimental Economics Laboratory (LEE) at Universitat Jaume I de Castellón —and particularly to José C. Pernías— for sharing their “summer spare computer capacity” with us. Their generosity helped us overcoming a tedious bottleneck in our initial computations. The second author acknowledges financial support through NSF Grant SES-0318208. We are solely responsible for all errors that might still remain.

[†] Department of Economics, University of Pennsylvania, McNeil Building / 3718 Locust Walk, Philadelphia, PA 19104-6297. Phone: 215-545-4287. E-mail: basaluzg@econ.upenn.edu

[‡] Department of Economics, University of Pennsylvania, McNeil Building / 3718 Locust Walk, Philadelphia, PA 19104-6297; and CEPR, London, UK. Phone: 215-898-1505. Fax: 215-573-2057. E-mail: miravete@ssc.upenn.edu; <http://www.ssc.upenn.edu/~miravete>

1 Introduction

There are very few contributions in the economics literature that compare the advantages of nonlinear tariffs relative to uniform pricing. According to Stole (2005, §6) this gap is due to the technical complexity of second-degree price discrimination, which impedes to obtain general results that allow us to evaluate the advantages of nonlinear over uniform pricing. However, important legislation such as the Robinson-Patman Act sets restrictions on the ability of firms to price discriminate, although as of today little is known about the welfare consequences of such long standing policies.¹ In the absence of clear theoretical results, the welfare comparisons among different pricing strategies remain ambiguous and this evaluation becomes an important open empirical question that has not been addressed either.² Our paper presents a framework where such comparisons as well as other policy evaluations are feasible while using the very limited tariff information generally available to economists.

The first theoretical comparison among different pricing mechanisms that we are aware of is the work of Spence (1977) who considers the problem of distributing a given amount of output among heterogeneous consumers (capacity pricing without exclusion). He proves the now well-known result that relative to uniform pricing, high valuation consumers purchase more and low valuation consumers purchase less with a nonlinear tariff. Spence also shows that the possibility of introducing discounts increases the monopolist's revenues but nothing is said about relative net welfare effects, neither the possibility of some consumers not purchasing at all is ever contemplated in his analysis. The work of Roberts (1979) first addresses the possibility of optimal exclusion of consumers in a model where consumer types include a vertical valuation dimension as well as income effects. However, no comparison of different pricing strategies are made either.³ Finally, Katz (1983) presents conditions where an increase in output associated to a change in a pricing strategy is a sufficient indicator for welfare improvement.

In all these models, a monopolist attempts to maximize profits by means of a nonlinear tariff while facing a heterogeneous customer base. Finding the optimal nonlinear tariff has long been stated by economists as a direct revelation mechanism. The solution of such standard problem consists of a pair of functions, $\{T(t), q(t)\}$, where a consumer of —possibly multidimensional— type t gets assigned

¹ Indeed, the 1936 Robinson-Patman Act amended Section 2a of the 1914 Clayton Act to make unlawful for a seller “to discriminate in price between different purchasers of commodities of like grade and quality” where substantial injury to competition may result. In practice this has meant important pricing restrictions for intermediate good markets (secondary line or alleged injury to rivals of the buyer receiving the discriminatory price). See O'Brien and Shaffer (1994) for further details.

² We should mention here the recent paper by Chan, Hall, and Rust (2004) where the comparison between a fully nonlinear tariff and a uniform price is briefly considered within a dynamic model of optimal inventory and sale decision.

³ This is a remarkable paper quite overlooked in the recent literature of nonlinear pricing. Using straightforward mathematical methods it establishes the basic results of nonlinear pricing models later developed by Armstrong (1996) and Rochet and Stole (2002). In particular Roberts (1979) proves that with two-dimensional types exclusion at the bottom is always optimal. He furthermore proves that the optimal tariff induces efficient pricing not only at the top as usual, but *also* at the bottom.

a consumption level $q(t)$ for an associated payment $T(t)$. Most empirical applications have only made use of the predictions of this general model regarding the behavior of $q(t)$ while ignoring the less data demanding use of the optimal tariff function $T(t)$.

This paper builds upon a recent line of research that makes use of the predictions of a nonlinear pricing model regarding the optimal tariff function $T(t) = T[q(t)]$. The information contained in the position and shape of tariffs has been recently used either to recover a single dimensional index of quality of products—as in Crawford and Shum (2001)—or alternatively, to identify the distribution of unobserved consumer heterogeneity both under monopoly—Miravete (2004)—and duopoly—Miravete and Röller (2004)—. In this paper we broaden this approach by presenting a functionally flexible formulation of the random participation nonlinear pricing model of Rochet and Stole (2002). We repeatedly solve this model numerically to identify the value of the structural parameters that generate the closest match to a large number of tariffs actually offered by monopoly carriers in the early U.S. cellular telephone industry between 1984 and 1988. The goal is to recover the parameters of the distribution of consumer types defined in a two-dimensional space: usage intensity of preferences (vertical heterogeneity), and individual-specific opportunity costs of participating in the market (horizontal heterogeneity).

Consumers may differ in several ways. A monopolist should, in principle, screen them with respect to as many dimensions as he can use to increase profits. Different type dimensions are defined by their induced price-independent shifts of individual demands. In telecommunications, for instance, tariffs may account for accumulated monthly usage, time of use (peak, shoulder, off-peak), distance (local, intrastate, interstate, international), frequency (unlimited calling and substantial discounts when calling family and friends), and identity of the network where calls are terminated (within own-network discounts and fixed/mobile interconnection surcharges) among others. Evidently these are not necessarily the only criteria to screen telephone customers. We should recall here the use of “creative pricing practices” such as short message services, news alert, stock market quotes, downloading of games, icons and tones, sending and receiving e-mails and pictures, internet access, and other services common to third generation mobile technology. A profit maximizing monopolist will make use of the dimensions that capture most of the heterogeneity of consumers. If consumers present an almost identical behavior in one given dimension the monopolist will most likely not screen them with respect to such dimension if implementing the screening device is costly. In an abstract setting where demands are linear we can distinguish, at most, two type dimensions summarized by the intercept and slope of demand, *i.e.*, the maximum willingness to pay for the product and the price responsiveness of each individual. Additional dimensions of consumer types will conform other features of the demand of consumers such as its degree of concavity or convexity or the number of related products that we have to account for. The added difficulty of multidimensional screening problems follows because it is generally impossible to order consumer preferences unequivocally. The identity of the consumer with the highest willingness to pay depends on the interaction of the different

type dimensions and is generally different for each possible price per unit of the product. This leads to optimal exclusion and bunching at the bottom —Armstrong (1996)— and non-monotonicity of optimal tariffs —Wilson (1993, §13-14)—, so that bunching for intermediate regions of the support of consumer types also occurs. The techniques to solve such problems are complicated and in practice difficult to be used in the empirical work.⁴ In theory almost any tariff behavior is possible under numerous regularity conditions. Thus, multidimensional screening models offer very little guidance to help identifying the structural parameters that lead to a tariff solution with some observed features.

The generalized single-dimensional screening model of Rochet and Stole (2002) —*RS* hereafter— is an interesting compromise between the standard single dimensional screening model of Mussa and Rosen (1978) and the general multidimensional screening model described in the previous paragraph and first discussed by Mirrlees (1971). We use a flexible formulation of the *RS* model to identify the structural parameters that best describe the demand and distribution of consumer heterogeneity. In Mussa and Rosen (1978) consumers valuations are ordered along a single dimension that also determines whether individuals participate in the market or not. If a consumer of type $t = 3$ finds valuable enough to participate in the market and consume, say $q(3) = 17$, then a consumer of type $t = 5$, with intrinsic higher valuation for every level of consumption possible, will *necessarily* participate in the market and consume $q(5) > 17$. This feature of the model —the well known single-crossing property— leads to a recursive variational problem that characterizes the optimal nonlinear tariff, a simplification that ensures that the incentive compatibility (*IC*) constraint is binding only upwards and can be enforced locally. Indeed, the optimal tariff is found by solving a first order nonlinear differential equation with a single boundary condition given by the reservation utility of the lowest active consumer type.⁵

In the *RS* model consumers differ in their valuation of the product, t , as well as in the value of their outside option associated to non-participation in the market, x . In this simplified framework, all additional sources of heterogeneity of consumers are summarized by the scalar x that enters additively into the utility function of consumers and is, by assumption, independently distributed from t . Thus, participation becomes endogenous because it depends on the alternatives available to each consumer. This feature of the model makes this framework particularly suitable to address competitive environments and will serve us to evaluate the potential benefits of divestiture of the monopolist into a symmetric duopoly. The existence of horizontal heterogeneity —the x dimension— reduces the ability of the monopolist to extract consumers' informational rents. Since participation is endogenous, it is optimal for the monopolist

⁴ See Rochet and Choné (1998) for the most general version available of this multidimensional screening model. Laffont, Maskin, and Rochet (1987) first provided with a closed form solution for such a problem in a model with two-dimensional types.

⁵ The slightly more general model of Maskin and Riley (1984) addresses screening consumers with respect to the different quantities that they purchase instead of the acquisition of a single unit of diverse quality. Footnote 5 of Rochet and Stole (2002) discusses the formal equivalence of both models although they interpret their own results from the perspective of quality discrimination among consumers with unit demands. Given our application to cellular pricing defined on airtime consumption, the interpretation of Maskin and Riley (1984) is more appropriate in our case.

to leave some additional surplus to a consumer of (vertical) type t in order to increase the likelihood that she participates in the market; a decision that depends on the realization of her (horizontal) type x . Consequently, the solution of the optimal nonlinear tariff with endogenous participation decision leads to a higher level of quality (quantity) for each type t relative to the solution of Mussa and Rosen (1978), which is achieved in the *RS* framework when the distribution of x becomes degenerate. Ignoring the possibility of horizontal heterogeneity may thus lead to an overestimation of the quantity underprovision needed to induce a full separation of different consumer types. However, making the participation decision endogenous comes at a significant cost. The variational problem that characterizes the optimal tariff is no longer recursive and instead it becomes a two-point boundary problem where a nonlinear second order differential equation provides the efficient quality (quantity) for the highest and lowest value in the support of t . Two-point boundary problems are difficult to solve in closed form. This explains why so few general features of the monopoly equilibria are presented in Rochet and Stole (2002, §4.1–4.2). In the present paper we use a flexible formulation and solve the optimal monopoly pricing numerically for different values of the structural parameters in order to match our model predictions to the actual tariffs offered by monopolists in several local markets in the early days of the U.S. cellular industry.

Why do we adopt the *RS* random participation framework to build our empirical equilibrium model? There are at least two compelling reasons to chose it over alternative models of single dimensional or multidimensional screening. First, the *RS* solution falls in between the solution of the single dimensional model of Maskin and Riley (1984) —*MR* hereafter— and the first best (*FB*) allocation where the monopolist is able to extract all the informational rents from each consumer type (first degree price discrimination). From an empirical point of view this feature of the *RS* model reduces the likelihood of misspecification. Assuming either full information or a single dimensional screening model would lead to inconsistent estimates if the participation decision is endogenous. In addition, estimating an *RS* based model does not rule out the possibility of concluding that the data is more consistent with either a single dimensional *MR* model or with the *FB* solution that arises in the absence of asymmetry of information. Second, although it is not the case for most of our data, nonlinear tariffs frequently include a monthly allowance of free minutes of airtime usage together with the payment of the fixed monthly fee. Wilson (1993, §6.4) shows that such an allowance is a constraint that needs to be imposed exogenously to characterize the equilibrium nonlinear tariff in the *MR* single dimensional screening model. Contrary to *MR*, the *RS* model endogenously explains this feature of the so-called “bucket tariffs.”

In this paper we present a method of recovering the structural parameters from a generalized screening model with endogenous participation decisions. We then use these parameters to address pricing issues that have been neglected in the literature and suggest potential policy applications of our methodology. We feel that our work contributes to the empirical literature on nonlinear pricing in at least two ways. First, we present an *RS* based model that can be solved numerically to conform actual tariffs.

Functional forms are chosen to allow flexible enough solutions, as well as to provide with economically intuitive interpretations of structural parameters. In estimating this model, we make use of the moment conditions that characterize the set of optional tariffs used to implement the optimal nonlinear tariff in practice. These moment conditions serve as the way to estimate a behavioral model where the different core parameters of the theoretical model are conditioned on observable demand and cost variables. Second, we use the structural estimates to evaluate the welfare performance of alternative pricing strategies such as uniform pricing, flat tariff, two-part tariff, and a Coasian tariff (marginal cost plus a fixed fee). An important effect of these different pricing mechanisms is to induce more or less participation of consumers. Thus, our framework allows us to conduct relevant policy evaluations such as to measure the costs *vs.* induced benefits of implementing a universal service requirement.

Our approach extends the work of Miravete (2004) and Miravete and Röller (2004) regarding the identification of structural parameters of a nonlinear pricing model from the shape and position of the tariff. As for the theoretical evaluations, we intend to fill the existing gap in the literature that has focused mostly on the characterization of equilibrium tariffs rather than on the desirability of price discrimination over uniform pricing. In relation to the policy evaluations, the universal service requirement is a commonly discussed issue in industries such as public utilities where nonlinear tariffs enjoy a widespread use. Policies aimed to achieve universal service are not exclusive of monopoly markets. Indeed, they have been actively used in promoting broadband access after the 1996 U.S. Telecommunications Act. Subsidies, however, may generate important inefficiencies depending on how they are financed.⁶ The evaluation carried out in this paper should be understood as an example of how to use the model to evaluate interesting alternative policies and illustrate the redistribution effects among consumers of diverse valuation that universal service policies may have. We certainly hope that the limited information required to conduct such an analysis and the accessible econometric methods employed will turn our approach appealing for policy makers and management of firms and utilities engaged in price discrimination strategies.

Our main results show that...

The paper is organized as follows. Section 2 intends to familiarize the reader with the institutions and regulations of the early U.S. cellular telephone industry. We also detail the sources of the data used in the empirical analysis of the paper. Section 3 presents our *RS* based model of nonlinear pricing with endogenous participation decisions. Section 4 presents a behavioral model linking the core parameters of the theoretical model to market specific demand and cost variables. Then, structural parameters are recovered by making use of the moment restrictions implied by firms' specific implementation of the

⁶ Crandall and Waverman (2000) estimate subsidies to ensure that schools, libraries, and rural health facilities have access to internet amount to \$2.65bn a year. Crandall, Hahn, Litan, and Wallsten (2004) argue in favor of removing retail price regulation of DSL providers as well as phasing out network sharing requirements as a far more effective strategy to induce investment in broadband capacity than direct subsidies. These authors also argue that broadband subsidies financed with taxation on other broadband services are particularly damaging to achieve a higher penetration rate in this market.

nonlinear tariff solution by means of a menu of self-selecting two-part tariffs. Section 5 evaluates the welfare of alternative pricing mechanisms relative to the unconstrained fully nonlinear tariff. Section 6 conducts the universal service policy evaluation. Finally, Section 7 concludes.

2 U.S. Cellular Monopoly Markets

By mid 1980s, the Federal Communications Commission (FCC) granted permission to create 305 non-overlapping cellular markets around U.S. standard metropolitan statistical areas (SMSA) to be served by two competing carriers. Cellular technology used low powered transmitters that exhausted the allocated bandwidth within small cells. A single conversation with the older high powered transmitters of car phones used a channel within a radius of about 75 miles. The FCC required that the low powered transmitters of the new cellular technology used these channels within a maximum radius of 30 miles.⁷

In 1981, the FCC set aside 50 MHz of spectrum in the 800 MHz band for cellular services. One of the two cellular channel blocks in each market—the B block or *wireline* license— was awarded to a local *wireline* carrier, while the A block —the *nonwireline* license— was initially awarded by comparative hearing to a carrier other than the local *wireline* incumbent. Licenses were awarded in ten tiers, from more to less populated markets, beginning in 1984. In general the *wireline* licensee offered the service first and enjoyed a temporary monopoly position until the *nonwireline* carrier entered the market, normally within six months of being awarded the license as required by the FCC. However, the administrative review process to award licenses among hundreds of contenders only based on technical issues and credibility of the announced investment commitments proved to be far more costly than initially expected. After awarding licenses for the thirty largest markets by means of this expensive and time consuming *beauty contest* —there were up to 579 contenders for a single license—, and while the application review of the second tier of 30 markets was on its way, rules were adopted to award the remaining *nonwireline* licenses through lotteries. Court appeals against the administrative awarding of the *nonwireline* licenses in the earlier tiers, and legal, technical, or managerial difficulties to start operating the lottery-awarded licenses in subsequent tiers led to a situation of temporary monopoly in many of the largest local cellular markets.

In this paper we use data from these monopoly markets. Data include detailed tariff information for about 45 *wireline* monopoly carriers between 1984 and 1988. The length of the monopoly phase in each market can be considered exogenous. Entry of the second firm depended mostly on court decisions regarding the contested administrative award of the *nonwireline* license by the FCC. Actually, entry of the second carrier always occurred soon after a firm court decision was made. Thus, predatory pricing incentives can safely be ruled out in this application.

Data include all tariff plans offered by each firm, their monthly subscription fee, rate per minute during peak hours, and the monthly allowance of free minutes, if any. We ignore off-peak pricing because at this early market cellular service is mostly targeting business customers: the handset was initially priced

⁷ Hausman (2002, §2.1) documents how the combination of cell splitting and sectorization increased capacity by a factor of 8 relative to the pre-existing, non-cellular, car telephone technology.

Table 1: Tariff Features

Option No.	Monthly Fee		Rate per Minute	
	Mean	Std.Dev.	Mean	Std.Dev.
<i>Markets with ONE option</i> (47 observations)				
1	32.12	(9.36)	0.36	(0.07)
<i>Markets with TWO options</i> (21 observations)				
1	17.37	(4.42)	0.56	(0.17)
2	38.70	(6.98)	0.35	(0.09)
<i>Markets with THREE options</i> (16 observations)				
1	3.25	(2.67)	0.64	(0.14)
2	15.62	(8.29)	0.41	(0.07)
3	34.01	(19.66)	0.33	(0.06)

Mean and standard deviations (between parentheses) of monthly fixed fees F_i and rate per minute p_i are measured in dollars.

at 3,000 dollars and peak pricing spanned over an 11 to 13 hour band at that time. In about one third of the markets, monopolists always offered the same tariff during all sample periods. Thus, we only include the earliest and latest tariff offered by each carrier during the monopoly phase of the cellular industry. In general, higher fixed monthly fees F_i go together with lower rates per minute p_i . Table 1 describes the features of the non-dominated tariff options offered by the firms of our sample. In general the addition of new tariff options to a firm's offering aims to expand his customer base by reducing the monthly fixed fee while increasing the tariff rate per minute of airtime. As we will see in our econometric analysis, this practice induce consumers with lower valuation to subscribe more frequently in those markets where firms offer more tariff options.

In addition, tariff data are complemented with market specific demand and cost information as well as an ownership indicator for each firm. This data will be used to conduct several correlation analysis of the parameters and welfare related effects recovered with our structural model with respect to observable characteristics of markets. Table 2 presents the descriptive statistics of our data. All money valued magnitudes are expressed in dollars of July 1986. Data definitions and sources are the following:

- Tariff information is reported by *Cellular Price and Marketing Letter*, Information Enterprises, various issues, 1984–1988. We also include the LATE indicator to identify whether the tariff information refers to the latest quarter of monopoly in each market.
- Socioeconomic and demographic data of each market comes from the 1989 *Statistical Abstracts of the United States*; U.S. Department of Commerce, Bureau of the Census, using the FCC Cellular Boundary Notices, 1982–1987, available in *The Cellular Market Data Book*, EMCI, Inc. Variables include the

Table 2: Descriptive Statistics

Variables	First Quarter		Last Quarter	
	Mean	Std.Dev.	Mean	Std.Dev.
PLANS	2.3333	1.4388	2.4651	1.3689
TCELLS	16.2308	16.6788	18.8372	21.1569
TIME	4.2308	2.1455	11.7209	3.8750
LATE	0.0000	0.0000	1.0000	0.0000
MKT-AGE	2.0000	2.9469	24.7209	10.4296
BUSINESS	41.2491	63.5971	51.7663	80.7145
COMMUTING	23.6490	3.4393	24.2024	3.7449
sdv(COMMUTING)	17.0208	3.0143	17.2975	3.4754
GROWTH	1.4026	0.9815	1.2651	0.9509
INCOME	38.1749	5.9610	38.7069	6.0354
sdv(INCOME)	31.7048	3.5367	31.9925	3.8184
EDUCATION	12.9533	0.5015	12.9660	0.4993
sdv(EDUCATION)	2.9365	0.2189	2.9533	0.2354
COVERAGE	0.0508	0.0207	0.0482	0.0184
POP-AGE	34.2225	2.0882	34.1069	2.0654
sdv(POP-AGE)	21.7781	1.0914	21.8292	1.2650
POPULATION	1.4678	1.6178	1.8207	2.1498
POVERTY	11.3897	3.1057	11.3581	2.9988
DENSITY	14.7833	14.2592	14.2689	12.5666
ENERGY	1.7517	0.4093	1.6724	0.3358
OPERATE	6.4542	1.6452	6.5012	1.5353
PRIME	10.8654	0.4962	9.0904	0.9318
RENT	16.3115	4.8599	16.2964	5.2371
WAGE	6.8753	1.5467	7.2801	1.6778
TEMPERATURE	52.1602	10.6482	61.1195	13.4410
RAIN	3.4863	1.9726	2.9279	2.0684
CRIME	6450.0983	1572.3222	7147.5182	2150.8539
VIOLENT	725.4234	331.7949	828.0634	372.8376
PROPERTY	5724.6750	1340.5812	6319.4548	1919.6731
SVCRIMES	0.1092	0.0325	0.1147	0.0369
REGULATED	0.4103	0.4983	0.4186	0.4992
AMERITECH	0.0513	0.2235	0.1163	0.3244
BELLATL	0.0256	0.1601	0.0000	0.0000
BELLSTH	0.2051	0.4091	0.1860	0.3937
CENDEL	0.0256	0.1601	0.0233	0.1525
CONTEL	0.1026	0.3074	0.0930	0.2939
GTE	0.0769	0.2700	0.0698	0.2578
NYNEX	0.0769	0.2700	0.0698	0.2578
PACTEL	0.0769	0.2700	0.1163	0.3244
SWBELL	0.0769	0.2700	0.0698	0.2578
USWEST	0.1026	0.3074	0.0930	0.2939
Observations	39		43	

All variables defined in the text.

number of months since the market started operating, MKT-AGE; thousands of high potential business establishments, BUSINESS;⁸ the average commuting time in minutes, COMMUTING; total population

⁸ BUSINESS refers to what it was considered at that time as highly potential customers by cellular industry experts: business service firms, health care, professional, and legal services, contract construction, transportation, finance, insurance, and real state.

of the SMSA in millions, POPULATION; average percent growth rate of population during the 1980s, GROWTH; median income in thousand of dollars, INCOME; percentage of households with income below the poverty level, POVERTY; median age of population, POP-AGE; and median number of years of education, EDUCATION. Variables $sdv(\cdot)$ indicate the standard within market standard deviation of the corresponding demographic.

- Industry cost indicators for each market are obtained from the Bureau of Labor Statistics; U.S. Department of Energy; *BOMA Experience Exchange Report: Income/Expense Analysis for Office Buildings*, various issues, 1985–1989; and *Cellular Price and Marketing Letter*, Information Enterprises, various issues, 1984–1988; and 1990 U.S. Census. They include the population density of the market (important for the deployment of antennas), DENSITY; the state average electricity rates in dollars per kilowatt/hour, ENERGY; one-period lagged prime lending rate, PRIME; an index of operating expenses per square foot of office space, OPERATE;⁹ an index of average monthly rent per square foot of office space in each market, RENT; and an index of average annual wages per employee for the cellular industry, WAGE.
- Weather data is available on the web at <http://cdiac.esd.ornl.org>, and includes average temperature and precipitation for 1,221 stations in the contiguous continental states plus those of Alaska.¹⁰ Data include the average quarterly temperature in Fahrenheit degrees recorded at the closest station to each market, TEMPERATURE; and the average quarterly precipitation in inches, RAIN.¹¹
- Crime information is obtained from the *Uniform Crime Report*, FBI, 1984–1988. We include the number of offenses per 100,000 inhabitants, CRIME; number of violent offenses per 100,000 inhabitants, VIOLENT; number of property offenses per 100,000 inhabitants, PROPERTY; and the percent share of violent crimes in each market, SVCRIMES.¹²
- Regulation was common in many markets as indicated by Shew (1994). The REGULATED dummy indicates that firms are required to get approval to offer new tariffs. The regulation regime was reported by the Cellular Telephone Industry Association in *State of the Cellular Industry*, 1992.

⁹ These expenses include cleaning, repair and maintenance, administrative costs, utilities, local taxes, security and ground services, office payroll, as well as other leasing expenses associated with running an office.

¹⁰ See Easterling, D.R., T.R. Karl, E.H. Mason, P.Y. Hughes, D.P. Bowman and R.C. Daniels, *United States Historical Climatology Network (U.S. HCN) Monthly Temperature and Precipitation Data*. ORNL/CDIAC-87, NDP-019/R3, 1996. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee.

¹¹ Climatology and location effects on the decision to subscribe to fixed local telephony has been documented by Crandall and Waverman (2000) and Riordan (2002, §2). The admittedly weak economic rationale is that people living in more inhospitable climates may spend more time indoors and may have a greater demand for fixed telephones and therefore lower demand for mobile telephony.

¹² There has been much speculation about the effect of crime as a driving force to subscription to cellular services. Indeed, cellular carriers at this early stage of the industry actively played this marketing strategy. See Murray (2002, p.212-213). Violent offenses include murder, non-negligent manslaughter, forcible rape, robbery, and aggravated assault. Property offenses include burglary, larceny-theft, motor vehicle theft, and arson.

- Largest shareholder information is available from the FCC. We identify the largest carriers: AMERITECH: Ameritech Mobile; BELLATL: Bell Atlantic Mobile; BELLSTH: BellSouth Mobility; CENTEL: Century Cellular; CONTEL: CONTEL Cellular; GTE: GTE Mobilnet; NYNEX: Nynex Mobile; PACTEL: PacTel Mobile Access; SWBELL: South West Bell; and USWEST: US West Cellular.

COVERAGE is an important variable that critically determines the dispersion of the distribution of horizontal heterogeneity, x . We define $COVERAGE = 1,300 \times TCELLS / (BUSINESS + 0.25 \times POPULATION)$, *i.e.*, we identify the potential market as that represented by a cellular telephone for each firm and families with an average number of four members. The total number of cells —antenna sites— of each firm was reported in several issues of *Cellular Business*. At this early stage of development of the industry, the number of subscribers were effectively constrained by the capacity of cellular carriers.¹³ This definition of market penetration is certainly arbitrary, but we feel that it is not unreasonable. We could have considered only the number business as the potential market. However, private use of cellular telephony also existed and cellular carriers targeted private users in their marketing campaigns as well. However, subscribing to a cellular service provider was far from common, and thus if we consider the total population as the potential market, coverage becomes minimal, implying necessarily that the distribution of the horizontal heterogeneity dimension, x , is almost degenerate. Our mix of BUSINESS and a fraction of POPULATION that approximates the number of families of each market appears to us as a reasonable definition of the potential market; helps identifying the parameters of our model; and thus highlights the methodological contribution of this paper and the value of its potential applications.

3 A Flexible Model of Nonlinear Pricing with Random Participation

This section presents an equilibrium model of monopolistic nonlinear tariff with optimal endogenous participation decisions of consumers. Our specification makes use of particular functional form assumptions to solve the RS model beyond their general but impractical characterization of equilibrium. Thus, our model does not add any theoretical result to those discussed by Rochet and Stole (2002); it rather allows us to implement empirically the RS model. Since all our functional form assumptions fulfill the requirements of the RS model the equilibrium tariff shares all the properties of the general solution. In this section we summarize the most relevant features of the equilibrium tariff and discuss the economic intuition behind our choice of functional forms.

¹³ Parker and Röller (1997, §4) report that one antenna could serve between 1,100 and 1,300 subscribers for the average use of cellular telephony at that time. Using a small sample of markets they also report that the correlation between the number of antennas and the number of subscribers exceeds 90%.

3.1 Basic Elements

A monopolist produces a good or service q at a constant marginal cost c and maximizes profits by designing the generally nonlinear and deterministic tariff $P(q)$:¹⁴

$$\pi(q) = P(q) - cq. \quad (1)$$

Consumers' preferences are indexed by a two-dimensional taste parameter (t, x) where t captures the vertical heterogeneity of consumers, *i.e.*, their intensity of preferences for the consumption of a given good; while x captures the horizontal heterogeneity of consumers, *i.e.*, an individual specific valuation of the outside option of not purchasing the product at all. The net utility from trade subtracts the outside opportunity cost from the indirect utility function u that is quadratic in consumption (leading to a linear demand for q):

$$v = tq - \frac{\gamma}{2}q^2 - P(q) - x, \quad (2)$$

where t and x are assumed to be independently distributed according to a Burr type XII distribution with parameter λ and an exponential distribution with parameter $1/\phi$, respectively:¹⁵

$$t \sim F(t) = 1 - \left[1 - \frac{t - \underline{t}}{\bar{t} - \underline{t}}\right]^{\frac{1}{\lambda}} \quad ; \quad \lambda \geq 0, \quad t \in \mathbb{T} = [\underline{t}, \bar{t}], \quad (3a)$$

$$x \sim G(x) = 1 - \exp\left(-\frac{x}{\phi}\right) \quad ; \quad \phi \geq 0, \quad x \in \mathbb{R}_+, \quad (3b)$$

with $c \leq \underline{t} \leq \bar{t}$. Finally, adding up (1) and (2) we obtain the joint surplus from trade:

$$S(q, t) = (t - c)q - \frac{\gamma}{2}q^2 - x, \quad (4)$$

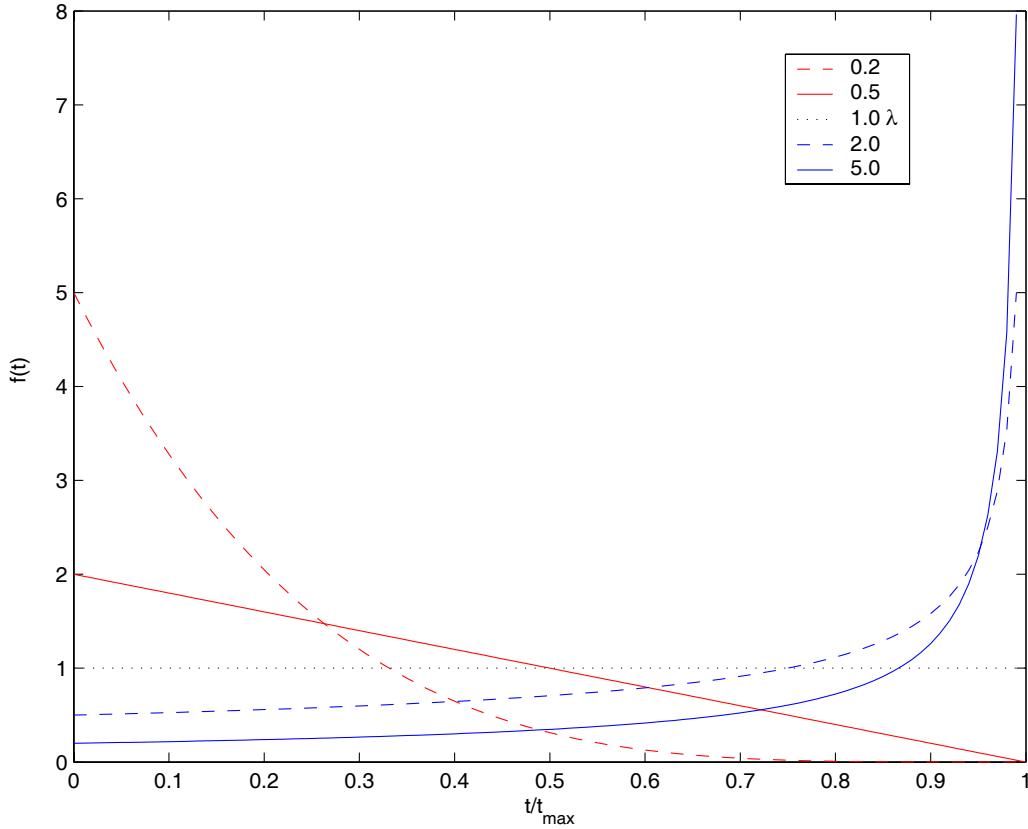
from where the *FB* solution is $q^{\text{fb}}(t) = (t - c)/\gamma$, *i.e.*, the purchase corresponding to the full information, first degree, price discrimination tariff.

The Burr type XII distribution of t in equation (3a) is a restricted beta distribution with parameters $(1, 1/\lambda)$. The interpretation of λ is appealing and intuitive. Different values of λ identify whether high

¹⁴ See Rochet and Stole (2002, footnote 15) for a detailed discussion in favor of deterministic contracts over randomization when consumers have to incur in transportation costs or when they can anonymously return to purchase repeatedly from the principal. Moreover, allocation to consumers using lotteries are not used in cellular telephone pricing.

¹⁵ Rochet and Stole (2002) also include a parameter σ to capture the importance of transportation costs. Thus, their distribution of the horizontal heterogeneity dimension is $G(x) = 1 - \exp(-x/\sigma\phi)$. Using only the tariff information available, we encountered that ϕ and σ are not independently identified. Therefore, we opted for normalizing $\sigma = 1$ and then recover the distribution of x for a common transportation cost across markets. This approach adopts what appears to us as a reasonable working assumption: not subscribing a cellular phone is assumed to be equally costly in all cellular markets studied but the distribution of consumers' valuations of not subscribing to the cellular monopolist may differ across markets.

Figure 1: $F(t)$ — Burr type XII distribution

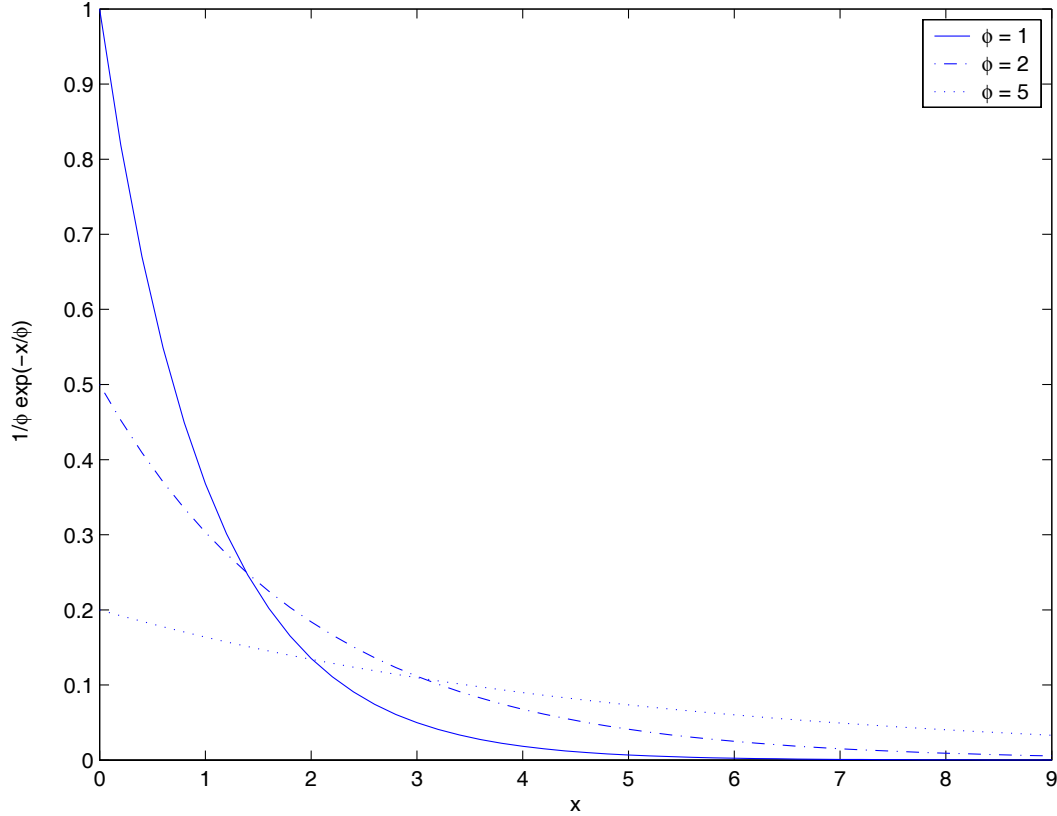


valuation consumers are more or less numerous than low valuation consumers. This is shown in Figure 1. If $\lambda = 1$ then t is uniformly distributed (*i.e.*, there are the same proportion of high and low valuation customers). If $\lambda > 1$, consumers are more concentrated around higher values of t (more numerous high valuation customers) and *vice versa* when $\lambda < 1$. If $\lambda = 0$, distribution (3a) becomes degenerate at $t = \underline{t}$, asymmetric information turns out to be irrelevant and homogeneous consumers can be efficiently priced by means of a single two-part tariff.¹⁶

Similarly, the exponential distribution (3a) is a good approximation to model the horizontal heterogeneity dimension, x . This exponential density is monotonically decreasing with a maximum at $x = 0$ and has an expected value of ϕ . This parameter increases when the average consumer has better outside options, perhaps offered by a competing firm in a horizontally differentiated industry. Thus, as it is shown in Figure 2, a small value of ϕ leads to a fast rate of decay of the probability density function of x . In terms of our model, it means that a small ϕ is associated with situations where consumer horizontal heterogeneity is also small. Thus, the probability associated to the event that consumers attach a large value to the outside option is very low. The contrary is true for large values of ϕ . The horizontal heterogeneity summarized

¹⁶ For further results and properties of the Burr type XII distribution see Johnson, Kotz, and Balakrishnan (1994, §12.4.5).

Figure 2: $G(x)$ — Exponential distribution



by x distinguishes the *RS* model from the standard single-dimensional *MR* screening model. If $\phi \rightarrow 0$ the horizontal heterogeneity disappears and we are left with a standard, single-dimensional, nonlinear pricing model. The role of x in the model is to capture the idea that consumers may have different valuations of their outside option of not purchasing q , independently of how their preferences are ranked with respect to this good through t . Thus, for large values of ϕ , it is likely that we find consumers with a high valuation of their outside option. These consumers will, most likely, not participate in the market. Thus, if the monopolist is to maximize profits, he must balance the alternative of extracting substantial informational rents from high t consumers with limited probability *vs.* increasing his customer base by lowering the markup charged to each consumer t . These opposite incentives explain most of the differences of the present model relative to the standard *MR* model, the most important of which is that in the *RS* model the monopolist has a reduced ability to extract informational rents, and thus, markups are lower than in the *MR* model.

Because we are unable to observe t and x , we decided to assume, —as *RS* did—, that type dimensions are independently distributed. If, for instance, these type dimensions were positively correlated, a higher valuation t goes together —although not perfectly— with a higher probability of participation in the market. The contrary would happen if t and x were negatively correlated. The strongest this negative

correlation is, the less able is the monopolist to extract informational rents from consumers, the further we are from the *MR* tariff, and the closer to the *FB* solution. Thus, if negative correlation exists, our model will overestimate ϕ , while this parameter will be underestimated if the correlation between x and t is positive. Furthermore, because of the existence of the individual specific outside option, x , the market share of the monopolist among consumer of type t is given by the following composition of distributions:

$$M(u, t) = \text{Prob}[t, x \leq u] = G(u)f(t) = \left[1 - \exp\left(-\frac{u}{\phi}\right)\right] \frac{1}{\lambda(\bar{t} - \underline{t})} \left(1 - \frac{t - \underline{t}}{\bar{t} - \underline{t}}\right)^{\frac{1}{\lambda} - 1}, \quad (5)$$

where $f(t) = F'(t)$ and $g(t) = G'(t)$ are the probability density functions associated to (3a) and (3b), respectively; $G(x)$ is log-concave —e.g., Karlin (1968, §1.5)—; and the following mean bound condition holds:

$$\lim_{x \rightarrow +\infty} xg(x) = \lim_{x \rightarrow +\infty} \frac{x}{\phi} \exp\left(-\frac{x}{\phi}\right) = 0. \quad (6)$$

Furthermore, the inverse hazard rate of M over u is nondecreasing in u :

$$H(u, t) = \frac{M(u, t)}{M_u(u, t)} = \phi \cdot \frac{1 - \exp(-u/\phi)}{\exp(-u/\phi)}, \quad \text{s.t.} \quad H_u(u, t) = \frac{1}{\exp(-u/\phi)} \geq 0. \quad (7)$$

3.2 Fully Nonlinear Tariff

Provided that an individual of type t reports her type truthfully as $u(t)$, which determines her decision to participate, the monopolist maximizes the expected profits designing the optimal direct revelation mechanism $\{P(t), q(t)\}_{t \in T}$ by maximizing the unconditional expected profits:

$$\int_{\underline{t}}^{\bar{t}} M(u(t), t) [S(q(t), t) - u(t)] dt, \quad (8)$$

with respect to $q(t)$ and $u(t)$ and subject to the participation or individual rationality and incentive compatibility constraints —*IR* and *IC*, respectively— i.e., $\dot{u}(t) = q(t)$ and $\dot{q}(t) \geq 0$ for all t . A piecewise-smooth function $q(t)$ is implementable by a tariff function $P(t)$ if and only if $q(t)$ is nondecreasing, $u(t)$ is absolutely continuous, and $\dot{u}(t) = q(t)$ at all continuity points of $q(t)$. This property of the *RS* model is common to the *MR* model.¹⁷ After substituting our specific functional form assumptions, this problem can be stated as:¹⁸

$$\max_{q(t), u(t)} \int_{\underline{t}}^{\bar{t}} \left[1 - \exp\left(-\frac{u(t)}{\phi}\right)\right] \left(\frac{\bar{t} - t}{\bar{t} - \underline{t}}\right)^{\frac{1}{\lambda} - 1} \left[(t - c)q(t) - \frac{\gamma}{2}q^2(t) - u(t)\right] dt, \quad (9a)$$

$$\dot{u}(t) = q(t) \geq 0, \quad (9b)$$

¹⁷ See Rochet and Stole (2002, Lemma 2), Tirole (1989, §3.5), and Wilson (1993, §6.2).

¹⁸ We denote by $\dot{u}(t)$ and $\ddot{u}(t)$ the first and second derivative of the utility function with respect to the horizontal type dimension t , respectively.

$$\dot{q}(t) \geq 0, \quad (9c)$$

$$u(t) = tq(t) - \frac{\gamma}{2}q^2(t) - P(t) \geq 0. \quad (9d)$$

As indicated in the introduction, this is not a recursive problem because of the existence of an individual specific outside option so that consumers will only participate when $u(t) \geq x$. It is therefore not possible to incorporate the IC constraints directly. In addition, the profit function is not separable in $u(t)$ and $q(t)$ and is nonlinear in $u(t)$, therefore turning impossible to have it integrated by parts. Rochet and Stole (2002) discuss at length the features of this equilibrium that critically depend on the dispersion of the distribution of t . Briefly, tariff features can be summarized as follows:

1. If the highest valuation \bar{t} is sufficiently large relative to \underline{t} the equilibrium tariff leads to bunching at the bottom, *i.e.*, the optimal tariff is such that all consumers on $t \in \mathbb{T}^\circ = [\underline{t}, t^\circ]$ purchase the same amount q° at $P^\circ = t^\circ q^\circ - \gamma(q^\circ)^2/2 - u(t^\circ)$. Furthermore, since $G(x)$ is log-concave, \mathbb{T}° is ensured to be a compact set. Pooling of different consumer types will never occur in any other region of the type space. Furthermore, in the upper interval of the type space $\mathbb{T} \setminus \mathbb{T}^\circ = [t^\circ, \bar{t}]$ the optimal utility profile solves the following second-order two-point boundary problem:¹⁹

$$0 = \frac{1}{\phi} \exp\left(-\frac{u(t)}{\phi}\right) (\bar{t} - t) \left[u(t) - \frac{\gamma}{2}\dot{u}^2(t)\right] + \left[1 - \exp\left(-\frac{u(t)}{\phi}\right)\right] (\bar{t} - t) [2 - \gamma\ddot{u}(t)] \quad (10a)$$

$$+ \frac{\lambda - 1}{\lambda} \left[1 - \exp\left(-\frac{u(t)}{\phi}\right)\right] [(t - c) - \gamma\dot{u}(t)],$$

$$\dot{u}(t) = q(t) \geq 0, \quad \text{for } t \geq t^\circ \quad \text{with } \dot{u}(t^\circ) = q^\circ, \quad (10b)$$

$$u(t) = tq(t) - \frac{\gamma}{2}q^2(t) - P(t) \geq 0, \quad \text{for } t \geq t^\circ \quad \text{with } u(t^\circ) = u^\circ, \quad (10c)$$

$$q(\bar{t}) = q^{\text{fb}}(\bar{t}) \geq 0. \quad (10d)$$

Thus, the efficiency at the top result holds so that $p(q(\bar{t})) = c$, while low valuation customers are pooled and possibly excluded in order to increase the informational rents extracted from higher valuation types. However, if the support of $F(t)$ is not sufficiently spread, all consumers are served, *i.e.*, $\underline{t} = t^\circ$ or $\mathbb{T}^\circ = \emptyset$, bunching does not exist and (10c) needs to be replaced by $q(\underline{t}) \geq q^{\text{fb}}(\underline{t}) \geq 0$, so that consumers are efficiently priced both at the *top* as usual, but also at the *bottom*. In short, if the monopolist finds profitable to serve the lowest type \underline{t} , he has to offer the most attractive price possible for this consumer to participate in the market for any realization of x . Thus, pricing at marginal cost becomes the optimal strategy.

¹⁹ For an advanced introduction to this topic see Ascher, Mattheij, and Russell (1988) or Keller (1968).

2. Provided that $G(x)$ is log-concave, full market coverage by a monopolist is never optimal regardless of whether the support of x is bounded or not.²⁰ This is an important feature of the model, —actually common to Armstrong (1996) and Roberts (1979)— because it allows us to evaluate the cost of promoting consumer participation beyond the monopolist’s optimal decision. In Section 6 we use the model to study the welfare gains induced by the *Universal Service* requirements commonly enforced in telecommunications.
3. The monopolist’s supply is such that $q(t) \in (q^{\text{mr}}(t), q^{\text{fb}}(t)]$ for all $t \in \mathbb{T} = [t, \bar{t}]$. Furthermore:

$$\lim_{\phi \rightarrow 0^+} q(t) = q^{\text{mr}}(t) ; \quad \forall t \in \mathbb{T}, \quad (11)$$

so that our estimates can continuously approximate to the *MR* solution and we can test whether the existence of the horizontal type x leads to a distortion —quantity underprovision for each type t — that is significantly smaller than in the *MR* case.

3.3 Qualifications

Before proceeding any further, we should point out few qualifications to our analysis. First, we are implicitly assuming that the monopolist solves a static problem every period. In our equilibrium approach, this rules out the possibility of addressing whether these *wireline* firms used the monopoly period to expand their customer base beyond the optimal static monopoly solution in order to deprive the entrant from the most valuable customers after having them subscribed to long term contracts. If this dynamic consideration were present, the actual markup will be below the optimal static markup and thus we will underestimate λ .²¹ As the optimal markup is less than the statically optimal one, it induces more participation than the statically optimal, and thus, ϕ would most likely be overestimated.

Second, the constant return to scale assumption of equation (1) ignores the possibility that capacity constraints may influence pricing decisions to allocate the cellular service among the highest valuation customers. If the actual tariff incorporated capacity pricing elements, our model would probably overestimate marginal costs and therefore λ , as the actual markup in the presence of capacity constraints would exceed that of static pricing with constant returns to scale. For the opposite reasons to those given in the previous paragraph, if capacity constraints are actually present our model would underestimate ϕ .

²⁰ This result is formally proven as Proposition 3 in Rochet and Stole (2002).

²¹ The inverse relation between the price markup and the hazard rate of the distribution is well documented in the agency theory literature. See for instance Laffont and Tirole (1993, §1.4-1.5) or Maskin and Riley (1984, §4). As the hazard rate of $F(t)$ is exactly $h_F(t) = 1/[\lambda(\bar{t} - t)]$, λ is directly related to the markup actually charged by the monopolist. Therefore, any downwards deviation from the optimal markup leads to downwards bias in the estimation of the true λ .

Third, we should consider the extent to which regulation might induce firms to deviate from profit maximization. It appears that this is not the case. Shew (1994, §4) indicates that the awarding of cellular licenses happened in a political environment that favored less and not more price regulation.²² In addition, cellular service—a luxury good at that time—was going to be provided competitively. Initially, only 12 states engaged in retail price regulation, some of which did not even disclosed the criteria for testing the reasonableness of cellular service prices. But even in those cases, regulation was vaguely enforced because even when at most state authorities required tariff filing, they let carriers to initially set their own price caps as regulatory bodies where uncertain to evaluate the costs of providing this new service.

However, the most important qualification to the *RS* model is to recognize that firms rarely offer the fully nonlinear tariff to their customers. Indeed, in our sample many firms only offer a single two-part tariff while others offer two, and a maximum of three self-selecting two-part tariffs. Therefore, instead of taking the number of tariff options offered as exogenously given, we will also incorporate the information contained in the number of tariff options actually offered by firms in each market. This will allow us to control for the distribution of a firm specific, commercialization cost, ζ , per tariff option offered to consumers. These commercialization costs include all non-observable incremental costs associated to the design, marketing, and advertising of each tariff option offered by firms. Since the incremental profits of adding an extra tariff option are positive but decreasing with the number of tariffs, a constant commercialization cost associated to each tariff option ensures that the monopolist find optimal to offer a menu with only a finite number of tariff options.²³ Next section incorporates the discrete choice of the number of tariff options within the *RS* framework and suggest how to estimate this model.

4 Empirical Analysis

This section addresses the empirical implementation of the *RS* model described above. We do this accounting for the actual number of two-part tariff options that different monopolists offer to their customers. We first describe the behavioral model relating the parameters of the *RS* model to market characteristics, therefore setting the exclusion of exogenous variables for each equation of the system. We then analyze the monopolist decision to screen consumers only through a menu with discrete number of tariff options, thus setting the exclusion restrictions involving endogenous variables. Finally, we present the estimates and discuss our results.

²² For instance, the 1980 Staggers Act had deregulated most railroad rates. Similarly, the 1978 Airline Deregulation Act mandated that the Civil Aeronautics Board terminated setting airfares beginning in 1983, although airfare competition was indeed allowed in May 1980. For further details, see Viscusi, Vernon, and Harrington (2000, §17).

²³ The approximate implementation of nonlinear tariffs by means of a finite menu of self-selecting two-part tariffs has only been addressed theoretically by Wilson (1993, §6.4). Recently Seim and Viard (2004) study how the number of tariff options offered change with the number of firms present in each market. Miravete (2004), within a *MR* framework, evaluates the foregone profits of not offering an additional tariff option and provides with an interval estimate of the magnitude of these commercialization costs.

4.1 Model Specification

A monopolist in market i at time τ observes the distribution of consumer preferences given by parameters $(\lambda_{i\tau}, \phi_{i\tau}, \bar{t}_{i\tau}, \bar{\gamma}_{i\tau})$; the slope of consumers' demand $\gamma_{i\tau}$; the marginal cost of production $c_{i\tau}$; and the cost of commercialization per tariff option $\zeta_{i\tau}$. We normalize $\bar{t}_{i\tau} = 0$, which implicitly assumes the systematic exclusion of low valuation types, a feature consistent with the low market penetration of the cellular telephone industry in the early age period of our sample. To complete the specification of the model, and make it suitable for econometric estimation we need to add some stochastic structure to this monopolist's maximization problem. Thus, each *core* parameter $\lambda_{i\tau}$, $\phi_{i\tau}$, $\bar{t}_{i\tau}$, $\gamma_{i\tau}$, $c_{i\tau}$, and $\zeta_{i\tau}$ is assumed to be a particular function of market specific characteristics observable to the econometrician, $\mathbf{Z}_{mi\tau}$, as well as other that remain unobservable and summarized by $\varepsilon_{i\tau}$, plus a vector of coefficients δ_m to be estimated for each equation so that we can write:

$$y_{mi\tau} = Y_m(\mathbf{Z}_{mi\tau}, \delta_m, \varepsilon_{mi\tau} | y_{-mi\tau}), \quad (12)$$

where $m = 1 \dots 6$, denotes the equation of each endogenous variable, *i.e.*, $y_{1i\tau} = \varphi_\lambda(\lambda_{i\tau})$, $y_{2i\tau} = \varphi_\phi(\phi_{i\tau})$, $y_{3i\tau} = \varphi_{\bar{t}}(\bar{t}_{i\tau})$, $y_{4i\tau} = \varphi_\gamma(\gamma_{i\tau})$, $y_{5i\tau} = \varphi_c(c_{i\tau})$, and $y_{6i\tau} = \varphi_\zeta(\zeta_{i\tau})$; and where functions $\varphi_m(\cdot)$ are possibly nonlinear transformations of the core parameters. Functions $Y_m(\cdot)$ specify different behavioral relationships between each core parameter and observable market characteristics. In our econometric specification we assume that all $Y_m(\cdot)$ are linear functions of market specific characteristics and possibly of other core parameters, $y_{-mi\tau}$. In particular, in Section 4.2 we postulate that realization of the marketing costs, ζ , determines the number of tariff options to be offered depending on the relative profitability of each market, which in turn is given by the magnitude of consumer's unobserved heterogeneity measured by λ , ϕ , \bar{t} , γ , and c .

Vector δ_m measures the effect of observable firm and market characteristics, $\mathbf{Z}_{i\tau}$,²⁴ and $\varepsilon_{i\tau}$ captures the effect of any relevant but unavailable information and/or any misspecification error (and whose components could potentially be correlated). These misspecification errors might indeed arise if demand is not linear, the industry does not operate under constant returns to scale, or more interestingly in our framework, if the true screening model includes more than two-dimensional types, or if within the family of two-dimensional type pricing mechanisms the effect of the outside option x enters non-additively into consumers' utility.

We postulate that demand related variables such as COMMUTING, BUSINESS, POPULATION, POP-AGE, EDUCATION, GROWTH, and INCOME enter into the equations of all demand related parameters of the model:

²⁴ $\mathbf{Z}_{i\tau}$ is an $N \times K$ matrix that includes N observations of K all exogenous market variables available to us and whose descriptive statistics are summarized in Table 2, while matrices $\mathbf{Z}_{mi\tau}$ in equation (12) indicates only those K_m exogenous variables that directly affect the behaviour of each core parameter.

λ , ϕ , \bar{t} , and γ . In addition, other regressors enter exclusively in one or few of the equations of some of these demand variables.

Since λ captures the degree of vertical heterogeneity among consumers valuations of cellular service, we also include the available market specific standard deviation of certain demographics (COMMUTING, POP-AGE, EDUCATION, and INCOME) to control for the dispersion of consumer valuations. We also include POVERTY(1-POVERTY) —the variance of a Bernoulli distributed variable— to control for the heterogeneity behind $F(t)$. All these variables are included exclusively in the λ equation.

Core parameter ϕ identifies the horizontal heterogeneity of consumers, their valuation of their outside option, and thus, their decision on whether they subscribe to the cellular telephone service or not. Only in this equation we include weather and crime variables —TEMPERATURE, RAIN, CRIME, and SVCRIMES— that have been either found to have a significant effect on subscription in other studies, or explicitly used in marketing campaigns in the early stage of development of the cellular industry to foster subscription to this new service.

The equation of marginal costs c includes cost related variables WAGE, ENERGY, OPERATE, RENT, PRIME, and DENSITY. We include the square of the latter variable because network deployment costs are high in low density areas (spread network with high cost per unit) and high density urban areas (congestion costs). We also include the number of antennae deployed, TCELLS, and its square to control for potential effects of the scale on the unit costs of production.

The specification of the maximum willingness to pay per minute, \bar{t} , includes all regressors of the equations of γ . Differentiating equation (2) with respect to q we obtain the demand function at a given marginal rate p , so that the maximum willingness to pay for a minute of airtime, \bar{t} can be identified by solving $q(\bar{t}) = q_{\max}$, i.e., $\bar{t} = b_v + \gamma \cdot q_{\max}$, so that all variables affecting γ will also affect \bar{t} . The maximum consumption level q_{\max} is set at 500 minutes a month.²⁵

Other variables enter most or all equations. Thus, LATE and MKT-AGE enter all behavioral equations (12). They capture both exogenous dynamic time effects on demand and cost variables and endogenous consumer experience and learning by doing costs effects. Regulation may affect not only the magnitude of markups that firms charge, therefore affecting λ , but also the decision to subscribe, ϕ , and the cost efficiency of firms, c through investment incentive distortions. However we expected that demand responsiveness re-

²⁵ The RS model, as well as the MR or any other model of screening assume a compact support for the distribution of consumer types. This leads to a finite maximum consumption level. A tariff, however is “open ended” and in principle we could compute the bills for telephone usage levels well above the common consumption patterns. We set $q_{\max} = 500$. Cellular telephone usage in this early market was far from our current habits. Hausman (2002) indicates that by 1993 the national average cellular telephone usage did not exceed 160 minutes a month. For any consumption level above 300 minutes, consumers are already in the range where the “last” tariff plan is in effect. We feel that 500 minutes a month is a reasonably high maximum consumption, and that setting much higher levels for q_{\max} will distort the account of expected profit and welfare.

mains independent of the tariff regulation regime set by authorities, and thus REGULATED is only excluded in the γ and \bar{t} equations.²⁶

Finally, we have to address the determinants of the commercialization costs ζ . Unfortunately, we do not have any information available regarding operation research or marketing costs. We thus include all variables affecting the marginal cost c plus firm specific fixed effects to control for the identity of the largest shareholder behind each carrier.

4.2 Screening with a Limited Number of Tariff Options

The vector of core parameters $(\lambda_{i\tau}, \phi_{i\tau}, \bar{t}_{i\tau}, \gamma_{i\tau}, c_{i\tau}, \zeta_{i\tau})$ fully characterizes market i at time τ . In each market, a monopolist could find the optimal fully nonlinear tariff $P(t) = P[q(t)]$ by solving the second-order two-point boundary problem (10a)-(10d). However, since offering tariff options is costly, in addition to solving (10a)-(10d) the monopolist has to decide how many tariff options to offer to his customers in order to maximize expected profits, *i.e.*, he has to choose the menu of n two-part tariffs that best approximate the fully nonlinear tariff solution within the set of n affine functions. A pair $\{A, b\}$ where A represents the fixed monthly fee and b the marginal tariff per unit of consumption fully describes any optional two-part tariff. Thus, the monopolist maximizes:

$$n_{i\tau} \in \arg \max_{\nu \in \mathbb{N}} \pi^* (\lambda_{i\tau}, \phi_{i\tau}, \bar{t}_{i\tau}, \gamma_{i\tau}, c_{i\tau} | \nu) - \nu \cdot \zeta_{i\tau}. \quad (13)$$

where $\pi^* (\cdot | \nu)$ denotes the expected profits of a monopolist that offers the ν self-selecting two-part tariffs that best implements the optimal fully nonlinear tariff option given by equations (10a)-(10d), that is:

$$\pi^* (\lambda_{i\tau}, \phi_{i\tau}, \bar{t}_{i\tau}, \gamma_{i\tau}, c_{i\tau} | \nu) = \max_{\substack{A_1 < \dots < A_\nu \\ b_1 > \dots > b_\nu}} \pi (A_1, \dots, A_\nu, b_1, \dots, b_\nu, \lambda_{i\tau}, \phi_{i\tau}, \bar{t}_{i\tau}, \gamma_{i\tau}, c_{i\tau}). \quad (14)$$

Contrary to other core parameters, we cannot obtain a point estimate of ζ . The reason is that only the number of tariff options is known, and this discrete realization only allows to recover at most the distribution of these marketing costs. However, we can think of this discrete choice of n as determined by the realization of the latent variable ζ exceeding different thresholds given by the magnitude of foregone

²⁶ Regulation might also drive the number of tariff options offered, in particular at the earlier periods as firms might have offered numerous options to preempt later delays in the approval of new tariff options. This practice has been documented by Shew (1994). The number of options offered determines the degree of concavity of the tariff schedule, and thus condition the value of λ and ϕ .

profits of not offering an additional tariff option.²⁷ Therefore for each market i and time period τ we have the following observation rule:

$$n = \mathbf{I} [\pi^*(\cdot | n) - \pi^*(\cdot | n - 1) > \zeta > \pi^*(\cdot | n + 1) - \pi^*(\cdot | n)], \quad (15)$$

and where $\pi^*(\cdot | 0)$ denotes the expected profits with uniform pricing.

Observe that the sequential profit maximization process described in (13)-(15) determines the features of the structural form to be estimated. First, the number of tariff options offered is determined by the magnitude of expected foregone profits relative to the commercialization costs. Thus core parameters that determine these foregone profits of not adding an additional tariff option, $(\lambda_{i\tau}, \phi_{i\tau}, \bar{t}_{i\tau}, \gamma_{i\tau}, c_{i\tau})$, need to enter the equation of the commercialization costs, $\zeta_{i\tau}$. Similarly, given some commercialization cost level, the monopolist offers more or less tariff options depending on how disperse the distribution of consumer valuations are. Thus, the value of core parameters $(\lambda_{i\tau}, \phi_{i\tau}, \bar{t}_{i\tau}, \gamma_{i\tau}, c_{i\tau})$ are conditioned by the number of tariff options offered by the monopolist. As a new tariff option is introduced, there is a lower fixed fee and marginal rate than before. This facilitates the participation of lower valuation customers (lowers ϕ), induces consumption by the highest consumer types (increasing \bar{t}), and reduces the proportion of high consumer types among active consumers (lowers λ). Finally, we furthermore interpret core parameters $(\lambda_{i\tau}, \phi_{i\tau}, \bar{t}_{i\tau}, \gamma_{i\tau}, c_{i\tau})$ as representing orthogonal dimensions of the RS model, thus characterizing different elements such as individuals' vertical intensity of preferences, their value of the outside option, their price sensitivity, and marginal cost of production of the monopolists that are assumed to be independent of each other. Therefore, these parameters do not interact with each other. All these considerations leads the following specification of the behavioral model (12):

$$y_m = \theta_{m6}\zeta + \mathbf{Z}_m\boldsymbol{\delta}_m + \varepsilon_m, \quad m = 1, \dots, 5, \quad (16a)$$

$$\zeta = \sum_{j=1}^5 \theta_{6j} \ln(y_j) + \mathbf{Z}_6\boldsymbol{\delta}_6 + \varepsilon_6. \quad (16b)$$

4.3 Estimation

How do we compute the values of these core parameters for each market in its first and last quarter under monopoly? Consider first the case of those (numerous) markets where only a two-part tariff is offered. We need to identify five core parameters of the RS model: λ , ϕ , \bar{t} , γ , and c . Evidently, a two-part tariff only provides with two pieces of information: the fixed fee A_1 and the rate per minute of airtime b_1 . The

²⁷ Wilson (1993, §8.3) shows that as long as IC is fulfilled and the distribution $F(t)$ is increasing hazard rate—and the Burr Type XII distribution fulfills this requirement as long as $\lambda > 0$ —, the incremental profits of adding a tariff option are positive but decreasing in the number of tariff options offered. Thus, the incremental profits of offering an additional tariff option will eventually fall short of any fixed costs of commercialization, therefore effectively limiting the optimal number of tariff options offered.

computed values of the parameters need to predict the actual monthly fee and marginal rate offered in each market:

$$\text{FEE}_{i\tau} : \frac{\partial \pi^* (\lambda_{i\tau}, \phi_{i\tau}, \bar{t}_{i\tau}, \gamma_{i\tau}, c_{i\tau} | \nu)}{\partial A_j} = 0, \quad j = 1, \dots, \nu, \quad (17a)$$

$$\text{RATE}_{i\tau} : \frac{\partial \pi^* (\lambda_{i\tau}, \phi_{i\tau}, \bar{t}_{i\tau}, \gamma_{i\tau}, c_{i\tau} | \nu)}{\partial b_j} = 0, \quad j = 1, \dots, \nu, \quad (17b)$$

In the case of a single two-part tariff we still need to add at least three more conditions to be able to compute the parameters of interest. A first condition that we use is to require that the model not only predicts the tariff properly, but also the market penetration. Thus, we require:

$$\int_{t^o}^{\bar{t}} \left[1 - \exp\left(-\frac{u(t)}{\phi}\right) \right] \frac{1}{\lambda(\bar{t} - t)} \left(\frac{\bar{t} - t}{\bar{t} - \underline{t}} \right)^{\frac{1}{\lambda} - 1} dt = \text{COVERAGE}_{i\tau}, \quad (18)$$

where $M(u, t)$ is the model prediction of market penetration as given by equation (5), and where COVERAGE is our available measure of market penetration in each SMSA.

The second condition that we use identifies the maximum type \bar{t} out of an arbitrarily chosen maximum consumption level. Thus:

$$\bar{t} = b_\nu + \gamma \cdot 500, \quad (19)$$

where $\nu = 1, 2, 3$ depending on whether the monopolist offers one, two or three optional tariffs.

Finally, we require that the average bill equals \$100.00 a month.²⁸ This number partially overcomes our lack of information about the distribution of individual consumption, thus generating meaningful economic predictions. We require that:

$$\int_{t^o}^{\bar{t}} P[q(t)] \left[1 - \exp\left(-\frac{u(t)}{\phi}\right) \right] \frac{1}{\lambda(\bar{t} - t)} \left(\frac{\bar{t} - t}{\bar{t} - \underline{t}} \right)^{\frac{1}{\lambda} - 1} dt = 100. \quad (20)$$

Each optional tariff adds two more conditions as those in (17a)-(17b), thus exceeding the number of core parameters to compute. We therefore minimize the square of the sum of the errors implicitly defined by (17a)-(20) in order to compute the value of the core parameters independently for each market by means of a simulated annealing process.²⁹ Solving this problem market by market is reasonable unless firms only

²⁸ The 2002 Semiannual Wireless Survey of the Cellular Telecommunications & Internet Association indicates that for 1988 the average monthly bill for cellular service reached its historical peak at \$98.02.

²⁹ The implementation of the simulated annealing algorithm is standard and follows Goffe, Ferrier, and Rogers (1994).

Table 3: Estimates

	λ	ϕ	\bar{t}	γ	c
CONSTANT	-4.8292 (0.80)	5.8799 (0.19)	-16.4070 (1.21)	-23.4861 (1.70)	4.7399 (0.09)
TIME	0.0443 (0.54)	0.7847 (2.73)	-0.1388 (0.87)	-0.1454 (0.90)	2.1295 (0.41)
MKT-AGE	-0.0937 (1.47)	-0.1282 (1.11)	-0.0555 (0.36)	-0.0551 (0.36)	-4.5424 (0.96)
LATE	0.2119 (1.19)	-0.4118 (1.15)	0.2841 (0.86)	0.2913 (0.88)	12.3446 (1.44)
REGULATED	-0.1101 (1.43)	0.7915 (2.18)			1.3252 (0.43)
BUSINESS	-0.0140 (0.16)	-0.0350 (0.07)	-0.0640 (0.28)	-0.0697 (0.29)	
POPULATION	-0.0397 (0.77)	0.0334 (0.07)	-0.0044 (0.03)	-0.0012 (0.01)	
INCOME	-0.7589 (1.40)	0.9791 (0.68)	-0.5423 (0.81)	-0.5544 (0.82)	
POP-AGE	0.2229 (0.25)	1.4303 (0.39)	0.4917 (0.52)	0.4708 (0.48)	
EDUCATION	2.9298 (1.11)	-13.6730 (1.46)	9.5530 (1.30)	9.9574 (1.34)	
COMMUTING	0.1276 (0.37)	2.7322 (1.74)	-0.3501 (0.41)	-0.3758 (0.43)	
GROWTH	0.0189 (0.38)	0.0283 (0.20)	-0.0932 (1.81)	-0.0963 (1.86)	
POVERTY	-0.8104 (0.21)				
TEMPERATURE		0.0982 (0.15)			
RAIN		-0.0446 (0.17)			
CRIME		0.2496 (0.51)			
SVCRISES		-0.6364 (1.35)			
TCELLS					-2.6425 (1.09)
WAGE					-7.2091 (0.80)
ENERGY					5.6726 (0.92)
OPERATE					0.9776 (0.18)
RENT					1.1047 (0.23)
PRIME					-0.7529 (0.17)
DENSITY					-1.3838 (0.89)
R^2	0.5505	0.6845	0.4705	0.4695	0.4415

Generalized least square estimates with absolute-value, heteroskedastic-consistent t-statistics reported between parentheses.

offered a nationwide tariff service. Table ??? shows that firms do not always offer the same tariff across markets.³⁰

Table 3 indicates...

Before conducting the econometric analysis, we have to rationalize the position and shape of the tariffs offered by each monopolist into some metrics. We use the *RS* model to generate such mapping. The values of the computed core parameters are reported in Table 4.

5 The Welfare of Alternative Constrained Pricing

If a monopolist has full information, *i.e.*, if he could observe the types of each individual consumer, the nonlinear pricing solution will be efficient. If the minimum consumer valuation exceeds the marginal cost of production, monopoly profits reach the maximum possible level and equal total welfare, while consumer surplus is zero. In an environment of full information, nonlinear pricing clearly dominates the

³⁰ Recovering structural parameters in a first stage to exploit the cross-market variation of demographics to explain their behavior is similar to the procedure used by Crawford and Shum (2004) in a discrete type screening model and no different from the recovering the mean utility level in the model of competition with differentiated products of Berry, Levinsohn, and Pakes (1995).

Table 4: Core Parameters of the RS Model

	<i>First Quarter</i>		<i>Last Quarter</i>		<i>Both Quarters</i>	
	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.
<i>ONE Option</i>						
λ	0.3061	0.0673	0.3114	0.0497	0.3086	0.0588
ϕ	697.7152	466.5313	752.4215	491.8656	724.4864	474.6539
\bar{t}	3.5142	0.9524	3.3508	0.7202	3.4342	0.8417
γ	0.0063	0.0018	0.0060	0.0013	0.0061	0.0016
c	0.0001	0.0004	0.0000	0.0000	0.0001	0.0003
Observations	24		23		47	
<i>TWO Options</i>						
λ	0.0688	0.0497	0.0884	0.0485	0.0791	0.0489
ϕ	84.1161	126.4811	95.7358	121.6028	90.2026	120.9456
\bar{t}	45.8497	64.4240	25.9405	53.8125	35.4211	58.4757
γ	0.0910	0.1290	0.0512	0.1177	0.0701	0.1171
c	0.0012	0.0013	0.0014	0.0012	0.0013	0.0012
Observations	10		11		21	
<i>THREE Options</i>						
λ	0.0146	0.0117	0.0173	0.0148	0.0163	0.0134
ϕ	17.5450	11.9923	33.3483	38.8856	27.4221	31.9003
\bar{t}	124.0571	153.1488	163.8782	154.0155	148.9453	149.8235
γ	0.2474	0.3063	0.3271	0.3081	0.2972	0.2997
c	0.0008	0.0012	0.0006	0.0009	0.0006	0.0010
Observations	6		10		16	

Empirical distribution of the computed core parameters.

welfare performance of linear pricing unless the uniform price is competitive and coincides with marginal cost c . However, beyond this very limited scenario, little is known about the relative performance of different pricing mechanisms to screen a population of heterogeneous consumers whose types remain private information.

Consider a standard nonlinear pricing mechanism *vs.* a uniform price strategy. Under very general conditions the optimal tariff leads to quantity discounts in order to separate large from small customers. Thus, large consumers will purchase more than under uniform pricing and the contrary will be true for small customers. In principle, the larger the fraction of high valuation consumers is, the closer is this nonlinear tariff to the efficient pricing solution. The caveat to this argument is that it might be optimal to exclude low valuation consumers altogether, and thus, the welfare comparison of a nonlinear tariff relative to a uniform price becomes ambiguous and subject to severe nonlinearities. The lack of robust theoretical predictions on this matter and the difficulty to compute general solutions of different pricing mechanisms turns this comparison into an essentially empirical question that has, so far, attracted little attention.³¹

³¹ Miravete (2004) studies the foregone profits and welfare of implementing a standard nonlinear tariffs by a monopolist only approximately by means of a variable number of self-selecting, optional, two-part tariffs. Miravete (2005) evaluates the relative performance of sequential pricing mechanisms relative to standard, nonlinear, monopoly pricing, where the monopolist may offer

In some circumstances, such as those envisioned by the Robinson-Patman Act, price discrimination is ruled out altogether. But in most cases, firms choose how to implement their pricing mechanism attending to screening costs considerations and common practices in the industry. In this section we compare the welfare, profits, and market penetration induced by a set of constrained pricing strategies with the most general nonlinear tariff computed in Section 3.2. The set of constrained tariffs includes: uniform pricing, flat tariff, the optimal two-part tariff, and a Coasian two-tariff where the price per additional unit equals the marginal cost of production.

5.1 Constrained Tariffs

5.2 Quantity Underprovision

Descriptive statistics comparing first best with RS and MR.

6 Policy Evaluations: Enforcing the Universal Service Requirement

The meaning of universal service in fixed telephony has changed quite substantially over the past century. Theodore Vail, CEO of AT&T, first used the term *universal service* in AT&T's 1907 Annual report to support the idea of a single nationally integrated telephone service. Universal service was the declared strategy of a dominant carrier to eliminate competitors with incompatible and not interconnected systems until the 1920s. In later decades, universal service meant a deliberate policy of underpricing local residential connections by overpricing long-distance calls in order to ease access to residential customers. To achieve this goal, business were also required to pay more for local connections than residents.³² Thus, universal service went together with cross-subsidization among the different line of services that ended with the divestiture of AT&T in 1984.³³

Although cellular telephony has not been subject to universal service policies, other competitive services such as broad-band have been targeted as essential services worth subsidizing to ensure most of the population access to internet and avoid the feared *digital divide*. As Crandall and Waverman (2000, §8) document, the 1996 U.S. Telecommunications Act ensures the subsidized access to internet for schools, libraries and rural health facilities at an estimated cost of \$2.65bn a year, far more than the traditional

either a menu of two-part tariffs or a menu of fully nonlinear options. Finally, Miravete and Röller (2004) evaluate the welfare and profit performance of nonlinear tariffs *vs.* two-part tariff, flat tariff, and uniform pricing in duopoly.

³² Muller (1997) discusses in detail the origin, history, and development of the universal service policy in fixed telephony in the U.S. and Crandall and Waverman (2000, §1) summarize recent developments of the universal service policy.

³³ For a definition of subsidy-free pricing see Faulhaber (1975). Riordan (2002, §4.1) briefly summarizes the discussion on the cross-subsidization generated by universal service policies in fixed telephony.

Table 5: Tariff Features: Welfare, Usage, and Market Penetration

	<i>Nonlinear</i>	<i>Flat</i>	<i>Linear</i>	<i>Two-Part</i>	<i>Coasian</i>
<i>ONE Option</i>					
MONTHLY FEE	— — —	142.3332	0	32.4322	142.2642
RATE PER MINUTE	— — —	0	0.4936	0.3715	0.0001
PROFITS	1.0000	0.9104	0.9520	0.9916	0.9104
WELFARE	0.7366	0.7167	0.7298	0.7362	0.7168
WELFARE (*)	4.3240	4.2075	4.2842	4.3217	4.2081
MARKET PENETRATION	0.0410	0.0341	0.0509	0.0404	0.0341
AIRTIME USAGE	263.4144	349.7056	202.1366	265.2019	349.6691
UNDERSUPPLY	0.4097	0.4526	0.3903	0.4062	0.4527
<i>TWO Options</i>					
MONTHLY FEE	— — —	197.6190	0	5.3577	197.5074
RATE PER MINUTE	— — —	0	2.3191	2.2400	0.0013
PROFITS	1.0000	0.8520	0.7957	0.8974	0.8521
WELFARE	0.7406	0.7022	0.6659	0.6936	0.7022
WELFARE (*)	18.9275	17.9460	17.0173	17.7244	17.9461
MARKET PENETRATION	0.1865	0.0898	0.2311	0.1931	0.0898
AIRTIME USAGE	52.2820	116.5023	30.9275	40.7882	116.4830
UNDERSUPPLY	0.3983	0.4275	0.2920	0.3218	0.4274
<i>THREE Options</i>					
MONTHLY FEE	— — —	47.8350	0	0.2173	7.8212
RATE PER MINUTE	— — —	0	2.3202	2.3067	0.0006
PROFITS	1.0000	0.8285	0.7555	0.7778	0.8286
WELFARE	0.7446	0.7032	0.6579	0.6638	0.7032
WELFARE (*)	3.6322	3.4303	3.2092	3.2380	3.4303
MARKET PENETRATION	0.1427	0.0675	0.1853	0.1789	0.0675
AIRTIME USAGE	12.7644	28.5787	6.8514	7.2523	28.5763
UNDERSUPPLY	0.3900	0.4133	0.2719	0.2779	0.4133

Values for the mean parameter estimates. Monthly fee and rate per minute are measured in dollars. Expected profits and welfare are measured in percentages (100%=1) relative to the fully nonlinear tariff case (ignoring marketing costs). WELFARE(*) indicates the money value in million of dollars for a market with mean potential customer base of about 400.000 customers. Market penetration is also measured as percentage of subscribers over potential consumers while airtime usage per subscriber is measured in minutes. Undersupply indicates the ratio of total minutes sold in a market under each tariff relative to the sales with the first best solution (no fee and rate equal to marginal cost).

support for universal service in fixed telephony. The analysis that follows shows the hypothetical effects of implementing a universal service requirement in the early stages of development of the cellular telephone industry while *simultaneously* for the unobserved heterogeneity of consumers as accounted for by the design of the nonlinear tariff offered by each monopolist.

Our monopoly model predicts optimal exclusion at the bottom. This feature of the model is also a very convenient because although market penetration of fixed telephony is over 90% it never reaches full market coverage. We then use the model to provide with two alternative ways to implement and measure the effects of the universal service requirement:

1. We first compute a balanced-budget pricing solution that maximizes market coverage while the monopolist makes non-negative profits, $\mathbb{S} \geq 0$. Thus, using the structural parameters from each tariff, we solve the following problem for each market and time:

$$\max_{q(t), u(t)} \int_{t^\circ}^{\bar{t}} \left[1 - \exp\left(-\frac{u(t)}{\phi}\right) \right] \frac{1}{\lambda(\bar{t} - t)} \left(\frac{\bar{t} - t}{\bar{t} - \underline{t}} \right)^{\frac{1}{\lambda} - 1} dt, \quad (21a)$$

$$\mathbb{S} \leq \int_{t^\circ}^{\bar{t}} \pi[q(t)] \left[1 - \exp\left(-\frac{u(t)}{\phi}\right) \right] \frac{1}{\lambda(\bar{t} - t)} \left(\frac{\bar{t} - t}{\bar{t} - \underline{t}} \right)^{\frac{1}{\lambda} - 1} dt, \quad (21b)$$

$$\dot{u}(t) = q(t) \geq 0, \quad (21c)$$

$$\dot{q}(t) \geq 0, \quad (21d)$$

$$u(t) = tq(t) - \frac{\gamma}{2}q^2(t) - P(t) \geq 0. \quad (21e)$$

This is the best way in which our single product model can capture the idea of the actual cross-subsidization that was induced by the rate structure sponsored by the *FCC*. In 1950 the *FCC* set the criteria to divide non-traffic sensitive fixed costs of the telephone network between interstate and intrastate services. According to Crandall and Waverman (2000, §1), long-distance cross-subsidized local calling only by the amount of 3% of the non-traffic sensitive, local-network costs but by mid-1970s, the subsidy represented already 20% of such costs. Provided the distribution of vertical and horizontal heterogeneity of consumers in each market, solving problem (21a)–(21e) we can predict by how much participation would increase if the monopolist is forced to break-even, *i.e.*, when $\mathbb{S} = 0$: $M_1^{US}(u, t^\circ)$. Alternatively we could also compute the market penetration for a given level of direct subsidy from the government, \mathbb{S} . We understand our break-even policy as a lower bound estimate of the effect of the universal service policy.

2. The maximum market penetration possible is to ensure that anybody willing to pay a positive price for telephone service actually subscribes to it. As for our model, we have to solve a modified version of (21a)–(21e) subject to the additional constraint that $P(t^\circ) = 0$ regardless of whether there is full market coverage, $t^\circ = \underline{t}$, or not. This provides us with the new market penetration $M_2^{US}(u, t^\circ)$ and the associated cost of providing this level of universal service, \mathbb{S} . We therefore view this second measure of market penetration as an upper bound estimate of the effect of universal service policy.³⁴

³⁴ In practice the *FCC* has set specific programs targeting low income customers to induce them to subscribe to local telephone service. Social tariff programs include “Lifeline,” a \$7 monthly subsidy for qualifying families —twice the federal subscriber line charge—; the “Link-Up America” program that subsidizes initial installation charges up to \$30 plus interest charges for up to twelve months for up to \$200 on deferred connection charges. The cost of these programs in 1998 amounted to \$464m (\$2.67 a year per access line) according to Crandall and Waverman (2000). For a very detailed description of these programs see Mitchell and Vogelsang (1991, §11).

Table 6: Average Effects of Universal Service Policy

	<i>Balanced</i>	<i>Free</i>
<i>ONE Option</i>		
WELFARE	0.8692	1.0000
WELFARE (*)	5.1024	5.8704
MARKET PENETRATION	0.1162	0.1018
AIRTIME USAGE	184.6214	259.0038
UNDERSUPPLY	0.8136	1.0002
<i>TWO Options</i>		
WELFARE	0.8202	1.0000
WELFARE (*)	20.9618	25.5556
MARKET PENETRATION	0.4605	0.3639
AIRTIME USAGE	42.5707	67.3072
UNDERSUPPLY	0.8008	1.0005
<i>THREE Options</i>		
WELFARE	0.8169	1.0000
WELFARE (*)	3.9848	4.8780
MARKET PENETRATION	0.3747	0.2909
AIRTIME USAGE	10.1210	16.0546
UNDERSUPPLY	0.8122	1.0003

Mean sample values across markets. Welfare and market penetration is reported as percentage of the profits of the fully nonlinear tariff solution. The rest of variables are defined as in Table 5.

In general, the existing studies document that income-targeted programs are not very effective in promoting subscription to fixed telephony.³⁵ However, all these studies take the tariff offered by the monopolist as given. Our structural approach allows us to recalculate what the optimal tariff of the monopolist would be in the presence of two alternative ways to implement the universal service policy.

7 Summary and Conclusions

³⁵ See Crandall and Waverman (2000, §6), Eriksson, Kaserman, and Mayo (1998), and Garbacz and Thompson (1997).

References

- AMEMIYA, T. (1978): "The Estimation of a Simultaneous-Equation Generalized Probit Model." *Econometrica*, 46, 1193–1205.
- ARMSTRONG, M. (1996): "Multiproduct Nonlinear Pricing." *Econometrica*, 64, 51–75.
- ASCHER, U. M., R. M. M. MATTHEIJ, AND R. D. RUSSELL (1988): *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*, (Classics in Applied Mathematics, Vol. 13). Englewood Cliffs, NJ: Prentice-Hall.
- BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): "Automobile Prices in Market Equilibrium." *Econometrica*, 63, 841–890.
- CHAN, H. M., G. HALL, AND J. RUST (2004): "Price Discrimination in the Steel Market." Mimeo, Charles River Associates, Yale University, and University of Maryland.
- CRANDALL, R. W., R. W. HAHN, R. E. LITAN, AND S. WALLSTEN (2004): "Universal Broadband Access: Implementing President Bush's Vision." Mimeo, AEI-Brookings Joint Center for Regulatory Studies, Working Paper # 04-01.
- CRANDALL, R. W. AND L. WAVERMAN (2000): *Who Pays for "Universal Service"? When Telephone Subsidies Become Transparent*. Washington, DC: Brookings Institution Press.
- CRAWFORD, G. S. AND M. SHUM (2001): "Empirical Modeling of Endogenous Quality Choice: The Case of Cable Television." Mimeo, Duke University and Johns Hopkins University.
- CRAWFORD, G. S. AND M. SHUM (2004): "Monopoly Quality Degradation in Cable Television." Mimeo, University of Arizona and Johns Hopkins University.
- ERIKSSON, R. C., D. L. KASERMAN, AND J. W. MAYO (1998): "Targeted and Untargeted Subsidy Schemes: Evidence from Post-Divestiture Efforts to Promote Universal Service." *Journal of Law and Economics*, 41, 477–502.
- FAULHABER, G. R. (1975): "Cross-Subsidization: Pricing in Public Enterprises." *American Economic Review*, 65, 966–977.
- GARBACZ, C. AND H. G. THOMPSON, JR (1997): "Assessing the Impact of FCC Lifeline and Link-Up Programs on Telephone Penetration." *Journal of Regulatory Economics*, 11, 67–78.
- GOFFE, W. L., G. D. FERRIER, AND J. ROGERS (1994): "Global Optimization of Statistical Functions with Simulated Annealing." *Journal of Econometrics*, 60, 65–99.
- HAUSMAN, J. A. (2002): "Mobile Telephone." In M. Cave, S. K. Majumdar, and I. Vogelsang (eds.): *Handbook of Telecommunications Economics*, Vol. I. Amsterdam, The Netherlands: North-Holland.
- JOHNSON, N. L., S. KOTZ, AND N. BALAKRISHNAN (1994): *Continuous Univariate Distributions*, Vol. 1, 2nd edn. New York, NY: John Wiley & Sons.
- KARLIN, S. (1968): *Total Positivity*, Vol. I. Stanford, CA: Stanford University Press.
- KATZ, M. L. (1983): "Non-Uniform Pricing, Output and Welfare under Monopoly." *Review of Economic Studies*, 50, 37–56.
- KELLER, H. B. (1968): *Numerical Methods for Two-Point Boundary-Value Problems*. Waltham, MA: Blaisdell.
- LAFFONT, J.-J., E. MASKIN, AND J.-C. ROCHET (1987): "Optimal Nonlinear Pricing with Two-Dimensional Characteristics." In T. Groves, R. Radner, and S. Reiter (eds.): *Information, Incentives, and Economic Mechanisms*. Minneapolis, MN: University of Minnesota Press.

- LAFFONT, J.-J. AND J. TIROLE (1993): *A Theory of Incentives in Procurement and Regulation*. Cambridge, MA: MIT Press.
- LEE, L.-F. (1981): "Simultaneous Equations Models with Discrete and Censored Dependent Variables." In C. F. Manski and D. L. McFadden (eds.): *Structural Analysis of Discrete Data with Applications*. MIT Press.
- MASKIN, E. AND J. RILEY (1984): "Monopoly with Incomplete Information." *RAND Journal of Economics*, 15, 171–196.
- MIRAVETE, E. J. (2004): "The Limited Gains From Complex Tariffs." Mimeo, University of Pennsylvania.
- MIRAVETE, E. J. (2005): "The Welfare Performance of Sequential Pricing Mechanisms." *International Economic Review*, 46, forthcoming.
- MIRAVETE, E. J. AND L.-H. RÖLLER (2004): "Competitive Nonlinear Pricing in Duopoly Equilibrium: The Early U.S. Cellular Telephone Industry." Mimeo, University of Pennsylvania and Wissenschaftszentrum Berlin für Sozialforschung (WZB).
- MIRRLEES, J. (1971): "An Exploration in the Theory of Optimum Income Taxation." *Review of Economic Studies*, 38, 175–208.
- MITCHELL, B. M. AND I. VOGELSANG (1991): *Telecommunication Pricing. Theory and Practice*. Cambridge, UK: Cambridge University Press.
- MULLER, M. L., JR (1997): *Universal Service*. Cambridge, MA: MIT-AEI.
- MURRAY, J. B., JR (2002): *Wireless Nation. The Frenzied Launch of the Cellular Revolution in America*. Cambridge, MA: Perseus Publishing.
- MUSSA, M. AND S. ROSEN (1978): "Monopoly and Product Quality." *Journal of Economic Theory*, 18, 301–317.
- O'BRIEN, D. P. AND G. SHAFFER (1994): "The Welfare Effects of Forbidding Discriminatory Discounts: A Secondary Line Analysis of Robinson-Patman." *Journal of Law, Economics, and Organization*, 10, 296–318.
- PARKER, P. M. AND L.-H. RÖLLER (1997): "Collusive Conduct in Duopolies: Multimarket Contact and Cross-Ownership in the Mobile Telephone Industry." *RAND Journal of Economics*, 28, 304–322.
- RIORDAN, M. (2002): "Universal Residential Telephone Services." In M. Cave, S. K. Majumdar, and I. Vogelsang (eds.): *Handbook of Telecommunications Economics*, Vol. I. Amsterdam, The Netherlands: North-Holland.
- ROBERTS, K. W. S. (1979): "Welfare Considerations of Nonlinear Pricing." *Economic Journal*, 89, 66–83.
- ROCHET, J.-C. AND P. CHONÉ (1998): "Ironing, Sweeping, and Multidimensional Screening." *Econometrica*, 66, 783–826.
- ROCHET, J.-C. AND L. A. STOLE (2002): "Nonlinear Pricing with Random Participation." *Review of Economic Studies*, 69, 277–311.
- SEIM, K. AND V. B. VIARD (2004): "The Effect of Entry and Market Structure on Cellular Pricing Tactics." Mimeo, Stanford University Graduate School of Business.
- SHEW, W. B. (1994): "Regulation, Competition, and Prices in the U.S. Cellular Telephone Industry." Mimeo, The American Enterprise Institute.
- SPENCE, M. (1977): "Nonlinear Prices and Welfare." *Journal of Public Economics*, 8, 1–18.
- STOLE, L. A. (2005): "Price Discrimination and Imperfect Competition." In M. Armstrong and R. H. Porter (eds.): *Handbook of Industrial Organization*, Vol. 3. New York, NY: North-Holland.

TIROLE, J. (1989): *Industrial Organization*. Cambridge, MA: MIT.

VISCUSI, W. K., J. M. VERNON, AND J. E. HARRINGTON, JR (2000): *Economics of Regulation and Antitrust*, 3rd edn. Cambridge, MA: MIT Press.

WILSON, R. B. (1993): *Nonlinear Pricing*. New York, NY: Oxford University Press.

Appendix

A Estimation of the Behavioral Model

Regardless of the particular behavioral model put forward by the econometrician, *core* parameters $\lambda_{i\tau}$, $\phi_{i\tau}$, $\bar{t}_{i\tau}$, $\gamma_{i\tau}$, $c_{i\tau}$, and $\zeta_{i\tau}$ fully capture the features of the model according to the present specification of the RS problem.... For convenience we define the $K \times K_m$ selection matrix \mathbf{J}_m that contains arrays of 0's and 1's such that:

$$\mathbf{Z}_{mi\tau} = \mathbf{Z}_{i\tau} \mathbf{J}_m. \quad (\text{A.1})$$

Since ζ is not observable, and only the number of tariff functions is available, (??)-(??) corresponds to a simultaneous equations model with discrete endogenous variables first considered by Amemiya (1978). The estimation of such model is straightforward and relies on imposing the identifying restrictions of substituting the reduced form parameters into the structural form (??)-(??).³⁶ Briefly, the estimation procedure is the following. We first estimate a set of reduced form equations of the endogenous variables on all K exogenous regressors:

$$\ln(\mathbf{y}_m) = \mathbf{Z}\beta_m + \mathbf{u}_m, \quad m = 1, \dots, 5, \quad (\text{A.2a})$$

$$\zeta = \mathbf{Z}\beta_6 + \mathbf{u}_6. \quad (\text{A.2b})$$

Next, making use of (??) and substituting (A.2a)-(A.2b) into (??)-(??) we obtain:

$$\ln(\mathbf{y}_m) = \theta_{m6} \mathbf{Z}\beta_6 + \mathbf{Z}\mathbf{J}_m \delta_m + (\varepsilon_m + \theta_{m6} \mathbf{u}_6), \quad m = 1, \dots, 5, \quad (\text{A.3a})$$

$$\zeta = \sum_{j=1}^5 \theta_{6j} \mathbf{Z}\beta_j + \mathbf{Z}\mathbf{J}_6 \delta_6 + \left(\varepsilon_6 + \sum_{j=1}^5 \theta_{6j} \mathbf{u}_j \right), \quad (\text{A.3b})$$

and thus, comparing the coefficients of (A.2a)-(A.2b) and (A.3a)-(A.3b) we obtain the set of identifying restrictions that allow us to estimate the structural parameters. Equations (A.2a) are estimated by ordinary least squares while (A.2b) is estimated as a maximum likelihood ordered probit model, thus producing consistent estimates $\tilde{\beta} = (\tilde{\beta}_1, \dots, \tilde{\beta}_6)'$ so that the identifying restrictions are evaluated as:

$$\tilde{\beta}_m = \theta_{m6} \tilde{\beta}_6 + \mathbf{J}_m \delta_m + \boldsymbol{\eta}_m, \quad m = 1, \dots, 5, \quad (\text{A.4a})$$

³⁶ In addition to Amemiya (1978), Lee (1981) further studies how to improve the asymptotic efficiency of this generalized least square estimator.

$$\tilde{\beta}_6 = \sum_{j=1}^5 \theta_{6j} \tilde{\beta}_j + \mathbf{J}_6 \delta_6 + \eta_6. \quad (\text{A.4b})$$

Since estimates $\tilde{\beta}$ are consistent and asymptotically normally distributed, the error terms $\eta = (\eta_1, \dots, \eta_6)$ are also asymptotically normally distributed:

$$\eta \sim N(0, \Sigma_\eta), \quad (\text{A.5a})$$

$$\Sigma_\eta = (\Theta \otimes \mathbf{I}_K) V(\tilde{\beta}) (\Theta \otimes \mathbf{I}_K)', \quad (\text{A.5b})$$

where $V(\tilde{\beta})$ is the covariance matrix of the reduced form estimates $\tilde{\beta}$ and Θ is the matrix of coefficients affecting the endogenous variables defined as:

$$\Theta = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & -\theta_{16} \\ 0 & 1 & 0 & 0 & 0 & -\theta_{26} \\ 0 & 0 & 1 & 0 & 0 & -\theta_{36} \\ 0 & 0 & 0 & 1 & 0 & -\theta_{46} \\ 0 & 0 & 0 & 0 & 1 & -\theta_{56} \\ -\theta_{61} & -\theta_{62} & -\theta_{63} & -\theta_{64} & -\theta_{65} & 1 \end{pmatrix}. \quad (\text{A.6})$$

The system of identifying restrictions (A.4a)-(A.4b) can be written in matrix form as:

$$\tilde{\beta} = D\theta^* + \eta, \quad (\text{A.7})$$

where:

$$D = \begin{pmatrix} \tilde{\beta}_6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{J}_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \tilde{\beta}_6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{J}_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \tilde{\beta}_6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{J}_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & \tilde{\beta}_6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{J}_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & \tilde{\beta}_6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{J}_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & \tilde{\beta}_1 & \tilde{\beta}_2 & \tilde{\beta}_3 & \tilde{\beta}_4 & \tilde{\beta}_5 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{J}_6 \end{pmatrix}, \quad (\text{A.8a})$$

$$\theta^* = (\theta_{16}, \theta_{26}, \theta_{36}, \theta_{46}, \theta_{56}, \theta_{61}, \theta_{62}, \theta_{63}, \theta_{64}, \theta_{65}, \delta'_1, \delta'_2, \delta'_3, \delta'_4, \delta'_5, \delta'_6)', \quad (\text{A.8b})$$

so that estimating (??) by generalized least squares to account for correlation among the elements of η provides with the system minimum distance estimator of the structural parameters:

$$\tilde{\theta}^* = (D' \Sigma_\eta^{-1} D)^{-1} D' \Sigma_\eta^{-1} \tilde{\beta}, \quad (\text{A.9a})$$

$$V(\tilde{\theta}^*) = \left(\boldsymbol{D}' \boldsymbol{\Sigma}_\eta^{-1} \boldsymbol{D} \right)^{-1}. \quad (\text{A.9b})$$