CARESS Working Paper 99–11

Quantity Discounts for Taste–Varying Consumers

Eugenio J. Miravete[†⋆]

September 17, 1999

Abstract

When a monopolist asks consumers to choose a particular nonlinear tariff option, consumers do not completely know their type. Their valuations of the good and/or optimal quantity purchases are only fully realized after the optional tariff has been subscribed. In order to characterize the menu of optimal nonlinear tariffs when consumers demands are stochastic, I assume that the distributions of the different components of consumers' types are log–concave to prove that the convolution distribution of these components is increasing hazard rate. This result, together with very weak assumptions on demand (common to standard nonlinear pricing), ensures that the continuum of optional nonlinear tariffs is characterized by quantity discounts. I test nonparametrically the model using data directly linked to consumer types from the 1986 Kentucky telephone tariff experiment. I show that the distribution of actual calls second order stochastically dominates the distribution of expected calls, which fully supports the suggested type–varying theoretical model. Finally, I analyze possible welfare effects of the introduction of optional tariffs and their relative expected profitability using the empirical distribution of consumer types in two local exchanges with differentiated calling patterns. The evidence suggests that a menu of optional two–part tariffs dominate any other pricing strategy. JEL: D42, D82, L96.

Keywords: Optional Nonlinear Pricing; Quantity Discounts; Convolution Distributions; Pólya Frequency Functions; Increasing Hazard Rate; Stochastic Dominance.

[†] Department of Economics. University of Pennsylvania, McNeil Building / 3718 Locust Walk, Philadelphia, PA 19104-6297; and CEPR, London, UK. Phone: 215-898-1505. Fax: 215-995-4186. E–mail: *miravete@econ.upenn.edu*

[⋆] I thank comments by seminar participants at Carlos III, Pompeu Fabra, Princeton, and Stern School of Business, as well as those attending the Worshop on Empirical Industrial Organization in Vigo. I am grateful to José Pernías for speeding up my simulation routines. Partial funding from the Ameritech Foundation through the Consortium for Research on Telecommunications Strategy and Policy is gratefully acknowledged.

1 Introduction

Consumers have to choose frequently among sets of class of services. For instance, telephone customers have to choose among different long distance plans offered by competing firms, or among different subscription contracts to the local telephone monopolist. Internet access providers also allow choosing among different connection plans depending on the expected usage of the network. Cable companies offer a variety of bundles of channels for monthly subscription at different rates and bundling discounts. Car rental rates depend on the duration of the lease, mileage, and/or fuel option chosen. Public transportation systems offer the possibility of advance purchase of passes of varied duration at different discount rates depending on the expected usage of the system. Banks ask their customers to select one among few checking and savings accounts depending on their average expected balance and number of monthly checks drawn. Also, insurance companies sell different policies conditional on the health, age, and/or deductible subscribed by the client. Finally, health clubs charge different monthly rates depending on registration fees related to the duration of the contract.

What do all these examples have in common? All these (and surely other) cases represent contractual situations where consumers choose a particular class of service, *i.e.*, a particular payment schedule, without knowing with certainty their future consumption level. Furthermore, for most of these examples (with the exception of insurance), consumers can be considered risk neutral without loss of generality given the low share of these expenditures on their average household income. Thus, the key issue is that telephone customers do not know their future local or long distance telephone usage. TV viewers do not know whether they will like future broadcasts or if they will have the time to enjoy it. Drivers do not know exactly the mileage of their planned trip, whether they will take a detour, or if they will face some contingency that force them to drive more or less than what they expected, *et cetera*. Therefore, consumers are not signing a contingent contract, and they are not committing to any particular purchase level in the future while the choice among alternatives does not embody any attempt to minimize risk.

All these situations are characterized by a two–stage decision process: first consumers decide which class of service they sign up for, and later, once their demand needs are known with certainty, they decide how much to buy from the firm, contingent on the rates of the tariff plan previously chosen. Companies, either because of reputation, repeated interaction with consumers, or legal restrictions, are not allowed to switch customers from one class of service to a different one, neither to take advantage of customers consumption decisions, or to favor them. Thus, firms can only profit from the stochastic dimension of consumers' demand through the design of the offered options.

But furthermore, most business applications of nonlinear pricing are characterized by quantity discounts, *i.e.*, unit price decreases with volume purchased by consumers. This feature is particularly convenient for natural monopolists and firms with important fixed costs. Charging a higher price per unit for the first units sold allows covering fixed costs, while discounts increase efficiency as large customers are priced closer to marginal costs. Fortunately, very general demand specifications and consumer taste distributions (adverse selection parameter) lead to optimal nonlinear price schedules characterized by quantity discounts. At least since the work of Maskin and Riley (1984), it is well known that quantity discounts are present if the distribution of types is increasing hazard rate, if consumers utility function satisfies the single–crossing property, and if some (not very restrictive) third derivatives of the utility function have the appropriate sign. Thus, general concave pricing mechanisms lead to quantity discounts. In fact, when the tariff function is concave, it can be implemented using a menu or continuum of self–selecting two–part tariffs whose marginal charge is decreasing with the volume purchased by the consumer [Faulhaber and Panzar (1977)].

Why cannot the above examples be addressed with the existing nonlinear pricing theory? The concept of self–selecting tariff has been wrongly used as synonym of optional tariffs. This is particularly true in many works dealing with pricing of telecommunications services. A common mistake present in all the related empirical literature is to neglect the existence of two stages and assume that consumers make purchases and choose among class of services simultaneously.[1] If this were the case, the only relevant information for consumers to make that decision would be known at the time of consumption, and therefore the "choice" of the corresponding self–selecting tariff plan would be exactly dual to the usage decision. Obviously, within this framework, there is no possibility of an *ex–post* "mistake" in the choice of the tariff plan.

The right approach should explicitly consider this two–stage nature of the problem: consumers first choose the tariff plan that better suit their needs given their expectation on future consumption level or expected valuation of the good or service. Later, once their needs or actual valuation are known, they decide how much to consume contingent on their previous tariff choice. The difficulty of this approach is that individual consumer' demands become stochastic since the expected consumption at the time of the tariff choice need not necessarily coincide with the purchase in the second stage of this game. Individual stochastic demands break the duality between consumption and choice of the corresponding self–selecting tariff. Consumers who chose different tariffs in stage 1 may end up paying and consuming the same at stage 2 if they receive demand shocks of opposite sign. Similarly, the same consumption level at stage 2 could be purchased at different rates depending on the disparate choice of tariffs of different consumers at stage 1. Thus, the optimal nonlinear tariff is no longer the lower envelope of a set of self–selecting two–part tariffs.

General treatments of nonlinear pricing when demand is stochastic are still not available.[2] This paper provides a characterization of the *ex–ante* nonlinear tariff (the one that considers two–stage decision problem) and relates it to the standard *ex–post* nonlinear tariff (where consumption and tariff choice are simultaneous). The key issue

[1] See for instance Hobson and Spady (1988), Kling and van der Ploeg (1990), MacKie–Mason and Lawson (1993), and Mitchel and Vogelsang (1991, §8).

[2] Only recently this topic has attracted some attention, although few and incomplete attempts to model optional tariffs have been carried out. See for instance Clay, Sibley, and Srinagesh (1992), Courty and Li (1998), and Miravete (1996 and 1997).

for these two pricing problems to be properly defined is that the hazard rate properties of the distributions of consumer types is preserved under convolution. This paper shows that this is the case under very general conditions, and furthermore that there might be some ordering of the hazard rates of the distributions used in each problem, *ex–ante vs. ex–post*, so that unit markups could also be compared.

The main goal of this paper is to study whether optional tariffs are characterized by quantity discounts. In all examples mentioned before, tariff plans with lower marginal charges were also those that include higher fixed fees. In particular, this paper aims to study whether Maskin and Riley's (1984) conditions still suffice to characterize optional tariffs with quantity discounts. If this is not the case, the paper explores which additional assumptions on demand and/or distribution behavior have to be made in order to ensure that optimal theoretical solutions are consistent with the common business practice. The paper shows that no additional assumptions on demand need to be made, and that in fact it suffices to assume that the distribution of taste components are log–concave to ensure that the corresponding optional tariffs are characterized by quantity discounts.

In order to deal with the stochastic nature of consumer demand, the suggested model assumes that consumers' types have two components: the *ex–ante* type θ_1, and the type shock θ_2. Together they define the *ex–post* type θ. The *ex–ante* type is always known by consumers, and in particular it is known prior to the choice of the class of service. This type dimension is private information and defines something similar to the average consumption level for each consumer (or expected valuation of the product). The type shock θ_2 represents deviations from the average consumption due to unpredictable events (or unexpected changes in valuation due to any general or individual circumstances). The type shock is different for each individual and remains private information to each consumer. The introduction of this second dimension captures the idea of changes in consumer preferences. This construction identifies the two sources of asymmetric information relevant at the choice and the usage decision stages. The monopolist will design each tariff option, and within each option the corresponding quantity discounts, to maximize his expected profits given the information set of consumers at each stage. The realization of θ_1 critically conditions the choice among tariffs, while the value of θ_2 together with the tariff plan chosen determines the actual level of usage in the second stage of the game.

Obviously, the application of this type–varying model is not limited to pricing problems, and it could be extended to any other mechanism design problem with stochastic demands. In particular, this paper is related to the regulation literature dealing with the optimality of linear contracts in the presence of cost randomness.[3] In terms of this literature the present paper analyzes whether the regulator is able to design a regulatory mechanism when there may be productivity shocks affecting the regulated firm, and whether it prefers to contract on the *ex–ante* expected costs of the regulated firm rather than on the actual *ex–post* costs.

[3] See Caillaud, Guesnerie, and Rey (1992), Laffont and Tirole (1986), and Rogerson (1988). See also Baron and Besanko's (1984) analysis of agents' type changes in a continuing regulatory relationship.

This type–varying setup rises many side issues that are intentionally neglected in the present paper to avoid unnecessary complexity and to limit the scope of the model in dealing with the effects of stochastic demands. I make two explicit simplifying assumptions here. First, the pricing game remains essentially static. Consumers first choose the optional tariff, and later decide how much to consume. I do not consider repeated versions of this game because it would require to model how informative is θ_1 with respect to θ, how are successive θ_2's correlated over time, and ultimately model the updating of future usage expectations, θ_1's. The closed–loop equilibrium tariff options of this richer model will be more difficult to characterize (if not impossible analytically), and this added complexity will however not help answering whether optional tariffs could be properly defined or not.[4]

The second simplifying assumption is that types remain single–dimensional. Thus θ "moves around" θ_1 depending on the magnitude of the type shock θ_2 and the single–dimensional definition of the *ex–post* type as a function of the *ex–ante* type and the shock. Additional dimensions should be considered if they address different attributes in the definition of consumers' utility functions, so that the monopolist can screen consumers different taste dimensions *simultaneously*. Nevertheless, in this paper I assume that consumers have heterogeneous preferences defined on just one single dimension. However, they have different knowledge about their preferences at each stage of the game. Consumers have a more or less intensive expected or actual valuation of the quantity consumed, but they do not take into account any other quality characteristic of the product. Thus, the stochastic nature of the problem allows the monopolist to screen *sequentially* each component of the *ex–post* type.[5]

An area where optional tariffs are prevalent is telecommunications. It is commonly reported in many telecommunications demand studies that telephone customers show a biased, even irrational, preference for flat tariff options. These studies conclude that by remaining on this option even when usage levels repeatedly do not justify it from a purely expense minimizing perspective, telephone customers are immensely increasing the cash revenues of telephone companies.[6] I exploit the unique data of the 1986 Kentucky telephone tariff experiment to test the empirical implications and make policy evaluations

---

[4] An additional but secondary reason is that I test this model with a short panel (three periods), and therefore learning effects cannot be consistently identified.

[5] Multidimensional screening is a complex but rich area itself. The main result of this literature is that type bunching is optimal due to a conflict between participation constraints and second order incentive compatibility conditions [Rochet and Choné (1998)]. My single–dimensional assumption also helps focusing the analysis on the stochastic feature of demand since multidimensional nonlinear pricing can only be solved explicitly for utility functions that are radial symmetric in type dimensions [Armstrong (1996), Wilson (1993, §12–14)], where monotonicity of the optimal tariff does not hold in general [Wilson (1995)]. Thus, within a multidimensional framework, it would not be possible to isolate whether the lack of monotonicity is due to the violation of any sufficient condition (increasing hazard rate of the distribution, single–crossing property, and sign restrictions of third derivatives of the utility function), or to the interactions among multiple type dimensions.

[6] See Hobson and Spady (1988), Kridel, Lehman, and Weisman (1993), and Srinagesh (1992). Train, Ben–Akiva, and Atherton (1989) use the same argument to explain the choice of tariff service to pay for domestic electricity consumption while Train, McFadden, and Ben–Akiva (1987) report that telephone customers switch options less frequently than expected from a pure cost minimization perspective.

using the suggested type–varying model. Since this data set includes direct instruments for θ and θ_1 (and therefore for θ_2 up to a monotone transformation), I can compute Anderson's (1996) nonparametric test of stochastic dominance to conclude that there is second order stochastic dominance of θ over θ_1 in the two exchanges of Kentucky (Bowling Green and Louisville) where the tariffs experiment took place. This evidence, consistent across time and demographic strata, strongly support the significance of the suggested type–varying model. The empirical analysis also shows that only in Louisville (the local exchange where the monopolist introduced optional tariffs) there is significant first order stochastic dominance (across time and demographic strata) of θ over θ_1. The advantage of using instruments directly linked to consumers' tastes, is that the empirical analysis is not subject to the common identification and misspecification of structural models dealing with asymmetric information issues.[7]

The paper is organized as follows. Section 2 presents the solution of the standard nonlinear pricing problem when the monopolist offers a continuum of *ex–post* self–selecting two–part tariffs, and studies whether they will be characterized by quantity discounts. Section 3 introduces optional nonlinear pricing, first through a menu of optional two–part tariffs, and later by means of nonlinear options. Section 4 proves that under very general assumptions the increasing hazard rate property of the distribution of consumers' private information parameter is preserved under convolution, and thus shows that *ex–post* pricing is well defined and generally characterized by quantity discounts when types include stochastically independent components. Section 5 presents evidence in favor of the taste-varying model using data from the 1986 Kentucky Local Telephone Tariff Experiment. Section 6 analyzes whether the monopolist and/or consumers prefer *ex–ante* to *ex–post* pricing, and empirically evaluates the welfare effects of the introduction of optional tariffs by using the kernel distribution of the observed θ, θ_1, and θ_2 in two local exchanges of Kentucky. Section 7 concludes.

2 Quantity Discounts in Nonlinear Pricing

This section briefly reviews the standard (*ex–post*) nonlinear pricing problem. I discuss the main assumptions of the pricing mechanism necessary to generate a separating Perfect–Bayesian Nash Equilibrium for the static game of incomplete information played by consumers and the monopolist. I also isolate sufficient constraints on demand and distribution of consumer's single–dimensional taste index so that screening of different types of consumers is achieved by means of quantity discounts. Because of the introduction of type changes in later sections of the paper, and in order to provide a reference framework to compare the solution of the optional nonlinear pricing mechanism given in Section 3,

[7] Empirical models as those of Ivaldi and Martimort (1994), Laffont, Ossard, and Vuong (1995), Miravete (1997), and Wolak (1996) identify the effects of asymmetric information through some structural restrictions and/or distribution assumptions, so that in fact it is difficult to acknowledge whether the estimated parameters actually isolate the effect of asymmetry of information or misspecification of the structural model.

I develop the mechanism using consumer's indirect instead of direct utility function, and marginal tariffs instead of quantities as the monopolist's control variable.[8]

I assume an environment where consumers' preference heterogeneity is captured by a single–dimensional index, θ. This taste indicator is private information for consumers while the monopolist only knows the population distribution of such index, $F(\theta)$. Given this informational constraint, the monopolist designs a fully nonlinear tariff to maximize his expected profits given the distribution of θ, by extracting consumer surplus in a different proportion depending of consumers' purchase levels. Thus, consumers are given incentives to self–select their purchase levels according to their preference intensity, θ.

The monopolist sells a single product x at a marginal charge p. Consumers' income is taken as *numeraire*. In addition, and for simplicity, I assume that there are no income effects for consumers or capacity constraints for the monopolist.[9] If consumers demand and the distribution of the private information parameter are properly behaved, the existence of quantity discounts is equivalent to the concavity of the tariff, which is just the lower envelope of the menu of self–selecting two–part tariffs.[10] This equivalence results guarantees the duality between the choice of any quantity under the nonlinear tariff and the unique corresponding choice of a marginal tariff and a fixed fee characterizing a hypothetical two–part tariff that leads to the same consumption level. Thus, for analytical convenience, I will assume that consumers choose the pair $\{A(x), p(x)\}$ instead of choosing x directly. The assumed indirect utility function net of fixed fee payment A is:

$$V(p, A, \theta) = v(p, \theta) - A = \int_{p}^{\infty} x(z, \theta) dz - A, \tag{1}$$

so that Roy's identity ensures that:

$$V_p(p, \theta, A) = v_p(p, \theta) = -x(p, \theta). \tag{2}$$

In order to characterize the optimal nonlinear schedule, some structure has to be imposed on the set of preferences as well as on the distribution of types. Focusing on demand, in order to ensure the existence of a separating equilibrium, it is necessary that

[8] The content of this section is however standard in the literature and a similar treatment using the direct utility function could be found in Tirole (1989, §3.5), as well as in Wilson (1993, §II) using the demand profile approach.

[9] See Oren, Smith, and Wilson (1985), Panzar and Sibley (1978), and Wilson (1993, §11) for treatments of capacity pricing. Goldmand, Leland, and Sibley (1984, §2), Ng and Weisser (1974), and Wilson (1993, §7) include income effects in nonlinear pricing modeling. Dealing with income effects does not add anything significant to the discussion of the present study while it complicates notation unnecessarily. Furthermore, the small share of local phone bills in consumers' incomes justifies the constant marginal utility of income assumption for the empirical application of this paper.

[10] If the number of two–part tariffs considered is discrete, this menu of two–part tariffs is equivalent to a declining block rate tariff that leads to a piecewise linear concave revenue function for the monopolist. See Brown and Sibley (1986, §4.4) and Faulhaber and Panzar (1977, §5).

consumers' demands do not cross so that consumers can be ranked by their preference intensity, θ. This is the well known *single–crossing property* (SCP).

ASSUMPTION 1: The indirect utility function is convex in price and increasing in the taste parameter. Thus $V_p(\cdot) = -x(\cdot) \leq 0$; $V_{pp}(\cdot) = -x_p(\cdot) > 0$; and $V_\theta(\cdot) > 0$; which implies $v_{p\theta}(\cdot) = -x_\theta(\cdot) < 0$ (SCP).

The monopolist's optimal tariff may be found by making use of the Revelation Principle in order to ensure that pricing mechanism is incentive compatible (IC) so that it induces consumers to self–select according to their true type. The IC constraint ensures that each consumer type keeps enough informational rents to consume according to their true preferences. This constraint of the monopolist's pricing problem is found by solving:

$$\theta \in \arg\max_{\theta'} \left[v(p(\theta'), \theta) - A(\theta') \right], \tag{3}$$

so that when choosing a particular two–part tariff each consumer chooses the one that maximizes her utility given the tariff schedule offered by the monopolist, *i.e.*, truthfull revelation of the *ex–post*, consumer private information. The solution to this problem, and thus the monopolist's IC constraint is:

$$A'(\theta) = -x(p(\theta), \theta)p'(\theta). \tag{4}$$

Next, by defining the rent of consumer type θ as:

$$V(\theta) = v(p(\theta), \theta) - A(\theta), \tag{5}$$

the Envelope Theorem allows us to represent the IC condition as follows:

$$V'(\theta) = v_\theta(p(\theta), \theta). \tag{6}$$

Finally, the monopolist also has to account for the individual rationality constraint (IR) since increases in fixed fees may induce some consumers not to participate in this market. Also for simplicity, I will assume that all consumers are served by the monopolist. It therefore suffices, because of monotonicity of the optimal marginal charge $\hat{p}(\theta)$, to ensure that the lowest valuation consumer $\underline{\theta}$ participates in the market:[11]

$$V(\underline{\theta}) \geq 0. \tag{7}$$

Hence, given (5) – (7) a monopolist with constant returns technology maximizes the following profit function:

$$\max_{p(\theta), V(\theta)} \int_{\Theta} \left[A(\theta) + (p(\theta) - c)\, x(p(\theta), \theta) - K \right] dF(\theta), \tag{8}$$

[11] The proof of this standard result is ommited. Because of SCP and IC, the chosen $\hat{p}(\theta)$ is nonincreasing in θ, and therefore it suffices that the IR constraint holds only at the lower bound of Θ_1.

where c is the constant marginal cost, and K the fixed cost of production for the monopolist. The solution of this problem (which is derived in Appendix 1) is a pair of functions $\{\hat{A}(\theta), \hat{p}(\theta)\}$ that relates each optimal two–part tariff offered by the monopolist to each consumer type θ:

$$\hat{p}(\theta) = c - \frac{1 - F(\theta)}{f(\theta)} \left[\frac{v_{p\theta}(p(\theta), \theta)}{v_{pp}(p(\theta), \theta)} \right], \tag{9}$$

$$\hat{A}(\theta) = v(\hat{p}(\theta), \theta) - \int_{\underline{\theta}}^{\theta} v_{\theta}(\hat{p}(z), z) dz. \tag{10}$$

Equation (9) presents the classical result that only the highest consumer type, $\overline{\theta}$ is efficiently priced. Since $v_{p\theta}(\cdot) < 0$ and $v_{pp}(\cdot) > 0$ by Assumption 1, all other types of consumers pay a marginal tariff higher than the marginal cost of production. The magnitude of the price distortion for each type θ therefore depends not only on the characteristics of demand (to induce self–selection), but also critically on the monopolist's knowledge of the population distribution of tastes. The spread of this distribution is related to the importance of the asymmetry of information between the monopolist and his customers regarding consumers' preferences. Therefore, conditional on the available information, the monopolist charges the optimal mark–up over marginal cost for each consumption level. The inverse relation between the hazard rate of the distribution of θ and the optimal marginal tariff capture the effect of the information about consumers' type distribution on the optimal monopoly pricing. The hazard rate of this distribution plays an important role in defining the magnitude of the price distortion (deviation from c) for each *ex–post* consumer type. Actually, in additon to SCP, I need to assume that $F(\theta)$ is increasing hazard rate (IHR) to ensure a separating equilibrium and avoid bunching of types at any given consumption or marginal tariff levels [Maskin and Riley (1984, §4)]. This property characterizes most common distributions used in economics, and the assumption should not be considered restrictive [Bagnoli and Bergstrom (1989)].

Definition 1: If a univariate random variable θ has density $f(\theta)$ and distribution function $F(\theta)$, then the ratio:

$$r(\theta) = \frac{f(\theta)}{1 - F(\theta)} \qquad \text{on} \quad \{\theta \in \Theta : F(\theta) < 1\}, \tag{11}$$

is called the hazard rate of either θ or $F(\theta)$.

Definition 2: A univariate random variable θ or its cumulative distribution function $F(\theta)$ are said to be increasing hazard rate if $r'(\theta) > 0$ on $\{\theta \in \Theta : F(\theta) < 1\}$.

Assumption 2: The *ex–post* preference index θ has a continuously differentiable probability density function $f(\theta) \geq 0$ on $\Theta = [\underline{\theta}, \overline{\theta}] \subseteq \Re$, such that the cumulative distribution function given by:

$$F(\theta) = \int_{\underline{\theta}}^{\theta} f(z)dz, \tag{12}$$

is absolutely continuous. Furthermore while θ remains private information for each consumer, $F(\theta)$ is common knowledge and IHR.

For the present model, the hazard rate function $r(\theta)$ indicates the probability that the monopolist faces a consumer of type $\theta' = \theta + d\theta$ given that he knows (*e.g.*, through any screening or sampling procedure) that the consumer's type is not smaller than θ. Therefore, if the distribution of types is IHR, the probability of finding a consumer within an arbitrarily small increase of the considered consumer type increases with the magnitude of the type of consumers, *i.e.*, the probability that one particular consumer falls in the next 1% segment of the market (in terms of purchase volume or willingness to pay for the good) is higher for the monopolist if he, in some way knows that one particular consumer belongs to the top 10% of the distribution of consumer valuations rather than to the top 30%.

Therefore, provided that the single crossing property holds, and the hazard rate of the distribution of *ex–post* types is increasing, the monopolist can discriminate among consumers by offering a continuum of self–selecting two–part tariffs that implement the optimal nonlinear pricing solution. Since each consumer type finds one and only one of these tariff plans maximize her utility, each two–part tariff is the optimal solution for only one *ex–post* consumer type, and therefore the equilibrium is ensured to be fully separating. As mentioned before, a sufficient condition for this continuum of two–part tariffs to be self–selecting is that its lower envelope be concave in consumption, *i.e.*, a tariff with quantity discounts. This is equivalent to the marginal tariff being decreasing in θ which requires:

$$\hat{p}'(\theta) = \frac{r'(\theta)}{r^2(\theta)} \frac{v_{p\theta}(p(\theta), \theta)}{v_{pp}(p(\theta), \theta)} - \frac{1}{r(\theta)} \frac{\partial}{\partial \theta} \left[\frac{v_{p\theta}(p(\theta), \theta)}{v_{pp}(p(\theta), \theta)} \right] \leq 0. \tag{13}$$

Since $F(\theta)$ has been assumed to be IHR, for the tariff to show quantity discounts it will suffice that the following condition, related to the second term on the right hand side of (13), and involving third derivatives of the indirect utility function, holds:

$$\frac{\partial}{\partial \theta} \left[\frac{v_{p\theta}(p(\theta), \theta)}{v_{pp}(p(\theta), \theta)} \right] \geq 0. \tag{14}$$

Any discrete version of the model will therefore consist of a menu of two–part tariffs where lower marginal rates $\hat{p}_1 > \hat{p}_2 > \ldots > \hat{p}_n$, are associated to higher fixed fees $\hat{A}_1 <$

$\hat{A}_2 < \ldots < \hat{A}_n$ –because of the IC equation–, and thus, these two–part tariffs will be self–selecting, and characterized by quantity discounts.[12]

3 Optional Nonlinear Tariffs

This section studies the design of optimal pricing mechanisms when consumer demand is stochastic. I analyze two cases: optional two–part tariffs and optional nonlinear tariffs. The sequential analysis of these two differentiated stages clearly points out the role of the distribution of each component of the type on the features of the tariff options. But first, before analyzing the design of tariff options, I have to define the *ex–post* type as a function of the *ex–ante* type and the type shock and how are they jointly distributed. The single–dimensional *ex–ante* type, *ex–post* type, and type shock are related as follows:

$$\theta = \theta_1 + \theta_2 \tag{15}$$

This definition of the preference intensity index is general. I discuss later in Section 4.3 how other functional forms only add unnecessary complexity to the problem without changing qualitative results. Next, as Assumption 2 defined the properties of the distribution of θ, I now have to specify the distribution of its component as well as their relation.

Assumption 3: The *ex–ante* preference index θ_1 and the type shock θ_2 have continuously differentiable probability density functions $f_i(\theta_i) \geq 0$, $i = 1, 2$, on $\Theta_1 = [\underline{\theta}_1, \overline{\theta}_1] \subseteq \Re$ and $\Theta_2 = [\underline{\theta}_2, \overline{\theta}_2] = [\underline{\theta} - \overline{\theta}_1, \overline{\theta} - \underline{\theta}_1] \subseteq \Re$ respectively, such that the cumulative distribution functions given by:

$$F_i(\theta_i) = \int_{\underline{\theta}_i}^{\theta_i} f_i(z)dz \qquad ; \qquad i = 1, 2 \tag{16}$$

are absolutely continuous. As in the standard case, θ_1 and θ_2 remain private information for each consumer while $F_1(\theta_1)$ and $F_2(\theta_2)$ are both common knowledge and IHR.

Assumption 4: The *ex–ante* type θ_1 and the type shock θ_2 are independent random variables.

This assumption is needed to solve the pricing problem explicitly, analyze how are the properties of the involved distributions related, characterize the properties of optional tariffs, and compare them to the standard nonlinear pricing solution. I also analyze in

[12] For inequality (14) to hold it suffices that $x(p, \theta)$ be concave in θ and that the price elasticity of demand be non–increasing in θ. These regularity conditions are rather technical and have little economic content, but they are assumed to hold in order to ensure a concave nonlinear schedule. When working the dual problem, using the direct utility function, this condition reduces to the SCP being not increasing in θ.

Section 4.3 how restrictive is this assumption and how would my results change if θ_1 and θ_2 were correlated.

3.1 Menu of Two–Part Tariffs

Two–part tariffs are widely used. They are effective in rising the necessary revenues to cover fixed costs and reduce the cost of monitoring customers' cumulative purchases as well as the incentives of arbitrage. From a marketing perspective optional two–part tariffs are attractive because of their simplicity. It is difficult to find examples where consumers have to choose among optional tariffs more complex than two–part tariffs per dimension of the product characteristic space. While more nonlinear optional tariffs will address better the differences among customers, and therefore will enhance welfare [Faulhaber and Panzar (1977, §4); Wilson (1993, §8.3)], consumers repeatedly show a biased preference for the simplest tariff option of the monopolist, or the competitor who offers the simplest tariff option.

Now consumers first choose an optional tariff characterized by a fixed payment A, and by a particular marginal tariff p. At the time of their choice, consumers are not fully aware of their preferences. They only know θ_1, and the distribution of θ_2. This means that consumers do not know how much will they consume when they choose the optional tariff plan. The choice of the tariff plan is final, and neither the monopolist can take advantage by switching consumers to a different plan, nor the consumer can request such a change in the interim between the tariff subscription and the consumption decision. If there is any tariff switching, it will only apply to future billing periods. Thus, given consumers' private information θ_1 and their expectation on type shocks, consumers choose the tariff plan that maximizes their expected net rent:

$$\theta_1 \in \arg\max_{\theta_1'} \int_{\Theta_2} \left[v(\tilde{p}(\theta_1'), \theta_1 + \theta_2) - \tilde{A}(\theta_1') \right] dF_2(\theta_2), \tag{17}$$

which given equation (15) and Assumption 4 leads to the following *ex–ante* IC constraint that applies to the choice of tariff options:

$$\tilde{A}'(\theta_1) = -E_2[x(\tilde{p}(\theta_1), \theta)]\tilde{p}'(\theta_1). \tag{18}$$

Next, the expected rent for a consumer of ex–ante type θ_1 is:

$$\tilde{V}(\theta_1) = E_2[v(\tilde{p}(\theta_1), \theta) - \tilde{A}(\theta_1)], \tag{19}$$

so that the IC constraint can be written as:

$$\tilde{V}'(\theta_1) = E_2[v_\theta(\tilde{p}(\theta_1), \theta)]. \tag{20}$$

Once the tariff option has been chosen, consumers learn their *ex–post* type through the realization of an individual type shock. Then their consumption level is decided, contingent on the previously chosen tariff plan $\{A(\theta_1), p(\theta_1)\}$. Since the consumption decision involves the choice of quantities, let $U[x]$ denote the corresponding direct utility function. The optimal consumption decision maximizes the actual rent given the tariff option:

$$\theta_2 \in \arg\max_{\theta_2'} U[x(\tilde{p}(\theta_1), \theta_1 + \theta_2')] - \tilde{A}(\theta_1) - \tilde{p}(\theta_1) x(\tilde{p}(\theta_1), \theta_1 + \theta_2'), \tag{21}$$

which leads to the *ex–post* IC constraint, *i.e.*, that consumers equate marginal utility to the marginal tariff at the consumption stage:

$$U'[x(\tilde{p}(\theta_1), \theta)] x_\theta(\tilde{p}(\theta_1), \theta) = \tilde{p}(\theta_1). \tag{22}$$

The monopolist has to consider two sets of participation constraints. To make the different solutions directly comparable, I assume that all households participate subscribing the service regardless of whether they later buy anything or not:

$$\tilde{V}(\underline{\theta}_1) \geq 0, \tag{23}$$

while *ex–post*, actual rents should be non–negative in order to consume and given the previous choice of service:

$$V(\theta_1, \underline{\theta}_2(\theta_1)) = v(\tilde{p}(\theta_1), \theta_1 + \underline{\theta}_2(\theta_1)) - \tilde{A}(\theta_1) \geq 0. \tag{24}$$

Thus, each *ex–ante* consumer type who chose the option $\{\tilde{A}(\theta_1), \tilde{p}(\theta_1)\}$ faces a different *ex–post* participation constraint. If the type shock is negative enough, $\theta_2 \leq \underline{\theta}_2(\theta_1)$, consumers with *ex–ante* type θ_1 do not to buy anything and thus the monopolist only gets the fixed fee $\tilde{A}(\theta_1)$ from them.[13] Then, given constraints (20) and (23) – (24), the monopolist maximizes the following profit function:

$$\max_{\tilde{p}(\theta_1), \tilde{V}(\theta_1)} \int_{\Theta_1} \left[\tilde{A}(\theta_1) + (\tilde{p}(\theta_1) - c) \int_{\underline{\theta}_2(\theta_1)}^{\overline{\theta}_2} x(\tilde{p}(\theta_1), \theta) dF_2(\theta_2) - K \right] dF_1(\theta_1), \tag{25}$$

The solution of this problem is again a pair of functions $\{\tilde{A}(\theta_1), \tilde{p}(\theta_1)\}$ that associates an optional two–part tariff offered by the monopolist to each consumer with *ex–ante* type θ_1:

$$\tilde{p}(\theta_1) = c - \frac{1 - F_1(\theta_1)}{f_1(\theta_1)} \left[\frac{E_2[v_{p\theta}(\tilde{p}(\theta_1), \theta) \mid \theta_2 \geq \underline{\theta}_2(\theta_1)]}{E_2[v_{pp}(\tilde{p}(\theta_1), \theta) \mid \theta_2 \geq \underline{\theta}_2(\theta_1)]} \right], \tag{26}$$

$$\tilde{A}(\theta_1) = E_2 \left[v(\tilde{p}(\theta_1), \theta) - \int_{\underline{\theta}_1}^{\theta_1} v_\theta(\tilde{p}(z), z + \theta_2) dz \;\middle|\; \theta_2 \geq \underline{\theta}_2(\theta_1) \right]. \tag{27}$$

---

[13] The *ex–ante* type dependent cut–off shock $\underline{\theta}_2(\theta_1)$ is uniquely defined in (24) for each θ_1 due to continuity of all functions involved and because $v_\theta(\cdot) > 0$ by Assumption 1.

This solution resembles that of the *ex–post* pricing very closely. With the exception of the ex–post participation constraint, the menu of optional two–part tariffs does not screen consumers with respect to their *ex–post* type, since θ_2 is integrated out in the monopolist's objective function. The monopolist just screens consumers with respect to their *ex–ante* type by offering them a menu of optional two–part tariffs that accounts for consumer differences before θ_2 is realized. The type shock only determines the amount that each consumer will purchase depending on the tariff option previously chosen. But the present pricing mechanism does not introduce any additional incentive to sort consumers who receive different shocks among those who chose a particular class of service. The results are still interesting. Denoting by $E_2^\star$ the conditional expectation with respect to the shock given that the ex–post participation constraints is fulfilled, and by differentiating (26) with respect to θ_1 we have:

$$\tilde{p}'(\theta_1) = \frac{r_1'(\theta_1)}{r_1^2(\theta_1)} \left[\frac{E_2^\star[v_{p\theta_1}(\tilde{p}(\theta_1), \theta)]}{E_2^\star[v_{pp}(\tilde{p}(\theta_1), \theta)]} \right] - \frac{1}{r_1(\theta_1)} \frac{\partial}{\partial \theta_1} \left[\frac{E_2^\star[v_{p\theta_1}(\tilde{p}(\theta_1), \theta)]}{E_2^\star[v_{pp}(\tilde{p}(\theta_1), \theta)]} \right] \leq 0. \tag{28}$$

Equation (28) shows that it is optimal to offer tariff options with lower marginal tariffs for future consumption if they are associated to higher actual fixed payments. Similarly (26) shows that only the higher *ex–ante* type is efficiently priced. All other *ex–ante* types are charged a positive mark–up directly related to the inverse of the hazard rate of the distribution of *ex–ante* types. These results are remarkable for two reasons. First, the *ex–ante* tariff is characterized by quantity discounts, which means that there is a concave, lower envelope function underlying the optional tariffs.

This concave function $\tilde{T}(\theta_1)$ is the mathematical lower envelope of the menu of two–part tariffs denoted by $\{\tilde{A}(\theta_1), \tilde{p}(\theta_1)\}$. But this function is not the tariff lower envelope in the traditional sense. For each *ex–ante* type θ_1 and tariff choice $\{\tilde{A}(\theta_1), \tilde{p}(\theta_1)\}$ there is a unique type shock $\theta_2 = \theta_2^*(\theta_1)$ such that total payments equal those of the lower envelope. We know that $\theta_2^*(\theta_1)$ is unique because the SCP requires that demand is increasing in the type, $x_\theta(\cdot) > 0$, and the marginal tariff $\tilde{p}(\theta_1)$ is given. Thus if consumers receive any other shock different from $\theta_2^*(\theta_1)$ they will move along the tariff option chosen and will always pay more under the chosen tariff regime than if the had "correctly" anticipated their future consumption, in which case they had moved along the lower envelope by choosing a different two–part tariff option. The monopolist may increase his revenues just because of this lock–in effect of the optional tariff if demand is stochastic. If we now repeat the analysis for other *ex–ante* types who chose different tariff options, we could easily check that the shape of the actual *ex–post* payment outlay is state–dependent, and that the payment outlay function is not ensured to be concave unless we unrealistically restrict the behavior of θ_2. However, since the distribution of θ_1 is IHR, the mathematical lower envelope $\tilde{T}(\theta_1)$ is still concave, and thus the optimal two–part tariff options are such that they lead to quantity discounts by offering a lower marginal rate associated to higher fixed fees.

The second reason that makes the result of equation (28) useful is because it shows that quantity discounts do not require any additional assumption on demand relative to those made for the *ex–post* case. The reduction of the marginal tariffs for class of services designed for high volumes of consumption is ensured because the distribution of *ex–ante* types is IHR, which suffices to ensure IC of ex–ante plans.

3.2 Menu of Nonlinear Tariffs

We can now deal with the general problem of a menu of nonlinear tariffs that also induce self–selection of consumers with respect to their type shocks. I will characterize this tariff using a constructive approach starting from the solution of the previous section. Each tariff option in Section 3.1 was a two–part tariff and thus, each consumer selected one among $\{\tilde{A}(\theta_1), \tilde{p}(\theta_1)\}$ depending on her *ex–ante* type θ_1 only but there was no further incentive for consumers to *ex–post* self–select according to the realization of θ_2. Two consumers with the same expected consumption or *ex–ante* valuation of the product would choose the same two–part tariff option. But they will later consume different amounts depending on their respective type shocks. Since both of them consume at the level where their marginal utility equated their marginal tariffs chosen in advance, the monopolist extracts some rent based on the consumer *ex–ante* type while consumers keep all their *ex–post* informational rent exclusively due to the learned type shock.

With nonlinear tariff options the monopolist also induces *ex–post* self–selection by means of quantity discounts and thus further reduces consumers' *ex–post* informational rents. Provided that each nonlinear tariff option is concave, they can also be represented by a continuum of self–selecting two–part tariffs. Hence, at stage 1, when consumers only know θ_1 they choose a nonlinear tariff option $\tilde{\tilde{T}}(\theta_1, \cdot)$, or alternatively a particular continuum of *ex–post* self–selecting two–part tariffs $\{\tilde{\tilde{A}}(\theta_1, \cdot), \tilde{\tilde{p}}(\theta_1, \cdot)\}$. Therefore, given consumers' private information θ_1, their expectation on type shocks, and their knowledge of the "shape" of the tariff options, they choose the tariff plan that maximizes their expected net rent.

General characterizations of the menu of nonlinear tariffs are difficult and cumbersome. However, the fact that types components are statistically independent from each other proves to be very useful in obtaining the characterization of this menu of nonlinear tariff options. Since the shock is independent of the *ex–ante* type, the mathematical lower envelope, $\tilde{T}(\theta_1)$, still capture the optimal incentive mechanism to screen consumers with respect to their *ex–ante* type dimension regardless of whether tariff options are two–part tariffs or more general nonlinear functions. Thus, $\tilde{\tilde{T}}(\theta_1, \theta_2)$ can be considered to be composed of two elements: one that screens consumers with respect to θ_1, represented by (26) – (27), and one that induces self–selection of *ex–post* types given the optimal tariff choices of each *ex–ante* type θ_1. This second component increases the revenue of the monopolist by reducing consumers informational rents exclusively related to θ_2. Obviously, if θ_1 and θ_2 were not independent it would be impossible to separate the origin of the rent extraction as screening for θ_1 should also account for the related distribution of θ_2. In order to characterize the optimal menu of nonlinear options $\tilde{\tilde{T}}(\theta_1, \theta_2)$, observe that since tariff options are nonlinear, the *ex–post* IC constraint of a consumer with *ex–ante* type θ_1 solves:

$$\theta_2 \in \arg\max_{\theta_2'} \left[v(\tilde{\tilde{p}}(\theta_1+\theta_2'), \theta_1+\theta_2) - \tilde{\tilde{A}}(\theta_1+\theta_2') \right], \tag{29}$$

i.e., a condition equivalent to (3). Observe also that since type components are independent each nonlinear option should be tangent only once to $\tilde{T}(\theta_1)$ as characterized by equations (26) – (27), so that $\{\tilde{A}(\theta_1), \tilde{p}(\theta_1)\}$ remains the optimal mechanism to screen the *ex–ante* type dimension θ_1 as the effect of θ_2 has been integrated out. Therefore, only one particular two–part tariff of each menu of *ex–post* nonlinear tariffs that characterizes each nonlinear option coincides with one of the optional two–part tariffs of the problem solved in the previous section. Thus total and marginal payments would be the same, as well as consumption, when the realized shock equals $\theta_2^*(\theta_1)$. This will not be true if *ex–ante* types and shocks were correlated because in such a case, the *ex–ante* choice of nonlinear tariff options will account for the effect of shocks on the marginal tariff paid *ex–post*, and thus $\tilde{p}(\theta_1)$ and $\tilde{\tilde{p}}(\theta_1, \theta_2^*(\theta_1))$ will generally differ. Under independence, the *ex–ante*, type specific, critical shock is implicitly defined by:

$$E_2[\tilde{\tilde{p}}(\theta_1+\theta_2)] = \tilde{\tilde{p}}(\theta_1+\theta_2^*(\theta_1)), \tag{30}$$

which reduces to:[14]

$$\frac{E_2^\star[v_{p\theta}(\tilde{\tilde{p}}(\theta_1+\theta_2), \theta_1+\theta_2)]}{E_2^\star[v_{pp}(\tilde{\tilde{p}}(\theta_1+\theta_2), \theta_1+\theta_2)]} = \frac{v_{p\theta}(\tilde{\tilde{p}}(\theta_1+\theta_2^*(\theta_1)), \theta_1+\theta_2^*(\theta_1))}{v_{pp}(\tilde{\tilde{p}}(\theta_1+\theta_2^*(\theta_1)), \theta_1+\theta_2^*(\theta_1))}. \tag{31}$$

The solution of the menu of nonlinear options builds upon the menu of optional two–part tariffs of Section 3.1. Taking the solution of the menu of optional two–part tariffs (26) – (27) as a boundary condition at $\theta_2^*(\theta_1)$, the optimal screening process with respect to θ_2 results in deviations of the *ex–post* marginal tariff and fixed fee payment relative to those of $\{\tilde{A}(\theta_1), \tilde{p}(\theta_1)\}$. Thus, let define:

$$\Delta\tilde{\tilde{p}}(\theta) = \tilde{\tilde{p}}(\theta_1+\theta_2) - \tilde{p}(\theta_1), \tag{32}$$

$$\Delta\tilde{\tilde{A}}(\theta) = \tilde{\tilde{A}}(\theta_1+\theta_2) - \tilde{A}(\theta_1). \tag{33}$$

The IC constraint (29) can be rewritten as follows:

$$\theta_2 \in \arg\max_{\theta_2'} \left[v(\tilde{p}(\theta_1) + \Delta\tilde{\tilde{p}}(\theta_1+\theta_2'), \theta) - v(\tilde{p}(\theta_1), \theta) - \Delta\tilde{\tilde{A}}(\theta_1+\theta_2') \right], \tag{34}$$

so that truthfull revelation of θ_2 maximizes the expected rent increase of switching from the "boundary, *ex–ante*, two–part tariff" to the corresponding *ex–post* two–part tariff on the chosen nonlinear option. Next, the rent is defined as:

$$\tilde{\tilde{V}}(\theta) = v(\tilde{\tilde{p}}(\theta), \theta) - \tilde{\tilde{A}}(\theta), \tag{35}$$

[14] Observe that uniqueness of $\theta_2^*(\theta_1)$ is ensured by monotonicity of the right hand side of this equation in θ as stated in (14).

and therefore, the IC constraint can be written as:

$$\tilde{\tilde{V}}'(\theta) = v_\theta(\tilde{\tilde{p}}(\theta), \theta). \tag{36}$$

In addition, there is another boundary constraint for this problem:

$$\Delta\tilde{\tilde{V}}(\theta_1 + \theta_2^*(\theta_1)) = 0, \tag{37}$$

so that $\tilde{\tilde{p}}(\theta_1+\theta_2^*(\theta_1)) = \tilde{p}(\theta_1)$, and thus each nonlinear tariff option is ensured to be tangent to $\tilde{T}(\theta_1)$ only once if $\tilde{\tilde{p}}(\theta)$ is monotone.

Given all these constraints the monopolist's problem solves, for each possible nonlinear option, the change in marginal rate that will maximize the increase in revenues from the corresponding "boundary two–part tariff" option:

$$\max_{\Delta\tilde{\tilde{p}},\Delta\tilde{\tilde{V}}} \int_{\Theta_2} \left[\tilde{\tilde{A}}(\theta) + \tilde{\tilde{p}}(\theta)x(\tilde{\tilde{p}}(\theta),\theta) - \tilde{A}(\theta_1) - \tilde{p}(\theta_1)x(\tilde{p}(\theta_1),\theta)\right] dF_2(\theta_2). \tag{38}$$

By pointwise maximization of this optimal control problem subject to (36), the first order necessary conditions are:

$$-\Delta\tilde{\tilde{p}}(\theta)v_{pp}(\tilde{\tilde{p}}(\theta),\theta)f_2(\theta_2) + \lambda_2(\theta_2)v_{p\theta}(\tilde{\tilde{p}}(\theta),\theta) = 0, \tag{39a}$$

$$f_2(\theta_2) = \lambda_2'(\theta_2), \tag{39b}$$

$$\lambda(\theta_2^*(\theta_1)) = 0. \tag{39c}$$

Observe that the transversality condition does not exist at $\theta_2^\star(\theta_1)$ since $V(\theta_2^\star(\theta_1)) > 0$ whenever $\theta_2^\star(\theta_1) > \underline{\theta}_2$, which ensures a unique tangency of each $\tilde{\tilde{T}}(\theta_1,\theta_2)$ to $\tilde{T}(\theta_1)$. Therefore:

$$\lambda_2(\theta_2) = \int_{\theta_2^*(\theta_1)}^{\theta_2} f_2(z)dz = F_2(\theta_2) - F_2(\theta_2^*(\theta_1)), \tag{40}$$

and thus, the optimal changes of the marginal tariff and fixed fee relative to the optimal two–part tariff option chosen by an *ex–ante* type θ_1 are:

$$\Delta\tilde{\tilde{p}}(\theta) = -\frac{F_2(\theta_2^*(\theta_1)) - F_2(\theta_2)}{f_2(\theta_2)} \left[\frac{v_{p\theta}(\tilde{\tilde{p}}(\theta),\theta)}{v_{pp}(\tilde{\tilde{p}}(\theta),\theta)}\right], \tag{41}$$

$$\Delta\tilde{\tilde{A}}(\theta) = v(\tilde{\tilde{p}}(\theta),\theta) - v(\tilde{p}(\theta_1),\theta) - \int_{\theta_2^*(\theta_1)}^{\theta_2} [v_\theta(\tilde{\tilde{p}}(\theta_1+z),\theta_1+z) - v_\theta(\tilde{p}(\theta_1),\theta_1+z)]dz. \tag{42}$$

These two equations in conjunction with (26) – (27) characterize the menu of optional nonlinear tariffs $\{\tilde{\tilde{A}}(\theta), \tilde{\tilde{p}}(\theta)\}$. Observe that equation (41) implies that consumers with *ex–ante* type θ_1 will face higher marginal charges than $\tilde{p}(\theta_1)$ if they receive a small shock $\theta_2 < \theta_2^*(\theta_1)$, but on the contrary, marginal tariffs would be smaller that $\tilde{p}(\theta_1)$ if $\theta_2 > \theta_2^*(\theta_1)$.

The final question that I have to address in this section is whether any further assumption is necessary to ensure that each nonlinear tariff $\tilde{\tilde{T}}(\theta_1, \theta_2)$ is concave, so that screening consumers with respect to their type shocks could also be achieved through quantity discounts. As Section 3.1 proved that $\tilde{T}(\theta_1)$ is concave, there only remains to analyze whether marginal tariffs $\tilde{\tilde{p}}(\theta)$ is decreasing in θ_2. For the purpose of interpretation, it is convenient to rewrite (41) as follows:

$$\Delta\tilde{\tilde{p}}(\theta) = \left[\frac{1}{f_2(\theta_2)}\frac{f_2(\theta_2^*(\theta_1))}{r_2(\theta_2^*(\theta_1))} - \frac{1}{r_2(\theta_2)}\right]\left[\frac{v_{p\theta}(\tilde{\tilde{p}}(\theta), \theta)}{v_{pp}(\tilde{\tilde{p}}(\theta), \theta)}\right]. \tag{43}$$

Thus, for each particular nonlinear option $\{\tilde{\tilde{A}}(\theta_1, \theta_2), \tilde{\tilde{p}}(\theta_1, \theta_2)\}$ to be concave it is required that:

$$\begin{aligned}\frac{\partial\Delta\tilde{\tilde{p}}(\theta)}{\partial\theta_2} &= \left[\frac{r_2'(\theta_2)}{r_2(\theta_2)} - \frac{f_2'(\theta_2)}{f_2^2(\theta_2)}\frac{f_2(\theta_2^*(\theta_1))}{r_2(\theta_2^*(\theta_1))}\right]\left[\frac{v_{p\theta}(\tilde{\tilde{p}}(\theta), \theta)}{v_{pp}(\tilde{\tilde{p}}(\theta), \theta)}\right] \\ &+ \left[\frac{1}{f_2(\theta_2)}\frac{f_2(\theta_2^*(\theta_1))}{r_2(\theta_2^*(\theta_1))} - \frac{1}{r_2(\theta_2)}\right]\frac{\partial}{\partial\theta_2}\left[\frac{v_{p\theta}(\tilde{\tilde{p}}(\theta), \theta)}{v_{pp}(\tilde{\tilde{p}}(\theta), \theta)}\right] \leq 0.\end{aligned} \tag{44}$$

The ratio $v_{p\theta}/v_{pp}$ is negative because the indirect utility function is convex in price, and because of the SCP as stated by Assumption 1, while its derivative with respect to θ_2 is ensured to be positive by equations (14) and (15). Hence, the concavity of the nonlinear tariff option critically depends on the signs of the terms between brackets. The proof of the following proposition shows that the distribution of θ_2 being IHR is no longer sufficient to ensure the concavity of the nonlinear tariff options.

Proposition 1: *Assumptions 1 and 2, do not suffice to ensure the concavity of each nonlinear option* $\tilde{\tilde{T}}(\theta_1, \theta_2)$.

Proof: The first term between brackets in equation (44) is ensured to be positive only if:

$$r_2'(\theta_2) > f_2'(\theta_2)\frac{1 - F_2(\theta_2^*(\theta_1))}{(1 - F_2(\theta_2))^2}. \tag{45}$$

Next, it is straightforward to show that the second term between brackets in equation (44) is negative only as long as $\theta_2 \leq \theta_2^*(\theta_1)$. ■

Therefore, if $f_2'(\theta_2) \leq 0$, it is also necessary that $\theta_2 \leq \theta_2^*(\theta_1)$ in order to ensure that each nonlinear tariff option is still concave. If $f_2'(\theta_2) > 0$ or if $\theta_2 > \theta_2^*(\theta_1)$, the IHR assumption, $r_2'(\theta_2) > 0$ does not suffice to ensure that each $\tilde{\tilde{T}}(\theta_1, \theta_2)$ is concave. Therefore we need the more restrictive assumption that the hazard rate of the distribution of the type shock is "increasing enough", in order to compensate the effect of large shocks or increasing density functions.[15]

If we just require that $r_2'(\theta_2) > 0$, we may find a curious asymmetric treatment of consumers with different *ex–ante* types. Nonlinear tariff options chosen by high *ex–ante* types are most likely concave. On the contrary, low *ex–ante* types choosing tariff options "designed" for low consumption levels would suffer important penalties (quantity premia) if they consume much more than what they expected.[16]

Hence, going from a menu of two–part tariffs to a menu of nonlinear tariffs only requires that the distribution of the shocks is increasing enough in hazard rate (IEHR) to ensure the existence of quantity discounts for every single nonlinear option. The basic setup remains however unchanged: the IHR property proves to be critical for the model to be well behaved. The following section studies how IHR properties are modified by the existence of stochastic components, so that the relative profitability of the *ex–ante vs. ex–post* pricing could be addressed.

4 Shocks, Convolutions, and Stochastic Dominance

The previous two sections have shown how to solve in isolation either the standard *ex–post* nonlinear pricing problem, or the more complex *ex–ante* optional nonlinear pricing problem. A most relevant question is whether these two solutions can be compared by the monopolist in order to choose the most profitable one in expected terms. Thus, these two problems have to be consistently defined so that the *ex–post* pricing solution accounts for all the statistical properties derived from the fact that θ is actually the result of the combination of θ_1 and θ_2.

In this section I focus on the relationship between properties of the distributions of the *ex–ante* type θ_1 and the type shock θ_2, and how are they related to the features of the distribution of the *ex–post* type θ. Showing that the IHR property of the distributions of the components of the type, $\{\theta_1, \theta_2\}$, is passed through to the distribution of the *ex–post* type, θ, is absolutely necessary to study the relationship between the features of the *ex–ante* optional and the *ex–post* standard nonlinear tariffs as well as to test the empirical implications of a model with type changes. I study these two issues in later sections of the paper.

[15] There are many cases where these conditions are fulfilled. One of such cases is when θ_2 is distributed as a standard unit beta distribution of the first kind with parameters $p = 1$ and $q = \lambda_2$. This standard unit Burr distribution of type 12 is IHR as long as $\lambda_2 > 0$, and the density function is always decreasing when $\lambda_2 > 1$. The hazard rate of this distribution varies from λ_2 when $\theta_2 = 0$ to ∞ when $\theta_2 = 1$. Thus, it is always possible to find a large enough value of λ_2 to ensure that the nonlinear tariff option is concave, even when $\theta_2 > \theta_2^*(\theta_1)$.

[16] Observe that to be consistent with equation (41), the marginal charges of the optional nonlinear tariffs should not exceed that of $\tilde{p}(\theta_1)$ whenever $\theta_2 > \theta_2^*(\theta_1)$.

The main result of this section is that if I focus on the class of log–concave probability density functions (all of which are IHR) to represent the distribution of the components of the *ex–post* type, then the distribution of the *ex–post* type is ensured to be log–concave, and therefore IHR. Thus, the two problems will be consistently defined, and both *ex–ante* and *ex–post* pricing will be characterized by quantity discounts as discussed in the previous sections. The rest of the analysis will be aimed to study the relative magnitude of these quantity discounts.

4.1 Preservation of IHR under Convolution

I should start this section by defining the distribution of the *ex–post* type in terms of the distribution of its components. As the following definition shows, the distribution of the *ex–post* type is just the Fourier convolution of its components.

DEFINITION 3: Let θ_1 and θ_2 be independent, univariate, random variables with cumulative distribution functions $F_i(\theta_i) : \Theta_i \to [0, 1]$, $i = 1, 2$. The cumulative distribution function of $\theta = \theta_1 + \theta_2$ is given by the convolution [Barlow and Proschan (1975, §4)]:

$$F(\theta) = \int_{\Theta_2} F_1(\theta - \theta_2) f_2(\theta_2) d\theta_2. \tag{46}$$

Therefore, given any arbitrary, but well behaved, distribution function for the *ex–ante* type (expected consumption or good valuation) and the type shock (unpredictable consumption or change in valuation after learning), it is always possible to identify the distribution of *ex–post* types up to a linear transformation. I will discuss in the next subsection whether the generality of my results is limited by the *ex–post* type definition (15) and Fourier convolution (46). The following property is extremely useful to identify the set of distributions that allow me to characterize the optimal optional tariffs.

DEFINITION 4: [Pečarič, Proschan, and Tong (1992, §13.5)]. Let $f_i(\theta_i) : \Theta_i \to [0, \infty)$. Then $f_i(\theta_i)$ is said to be log–concave if:

$$f_i[\alpha\theta_i + (1 - \alpha)\theta_i'] \geq [f_i(\theta_i)]^\alpha [f_i(\theta_i')]^{(1-\alpha)} \qquad \forall \theta_i, \theta_i' \in \Re, \text{ and } \forall \alpha \in [0, 1]. \tag{47}$$

I am interested in showing that if both the distribution of *ex–ante* types and the distribution of shocks are IHR, then the distribution of the *ex–post* type has to be IHR necessarily. The preservation of the IHR property under convolution is a direct consequence of the family of IHR distributions being totally positive functions. However, it is not necessary to deal with total positivity in general to show the main results of this section. I will therefore limit the scope of the analysis to a subclass of Pólya frequency functions.

DEFINITION 5: A function $g(z)$ is a Pólya frequency function of order 2 (PF_2) if $\forall x_1, x_2 \in X \subseteq \Re$ and $\forall y_1, y_2 \in Y \subseteq \Re$, such that $x_1 < x_2$ and $y_1 < y_2$, the following condition holds:

$$\begin{vmatrix} g(x_1 - y_1) & g(x_1 - y_2) \\ g(x_2 - y_1) & g(x_2 - y_2) \end{vmatrix} \geq 0. \tag{48}$$

Pólya frequency functions are just totally positive functions of order 2 in translation. Hazard rate properties of the distribution functions are easily characterized by log–concavity of the distribution functions while convolution properties are derived from the fact that IHR distributions are totally positive functions. The following lemma shows the equivalence between PF_2 and log–concave functions.[17]

LEMMA 1: *A twice continuously differentiable function $g(z)$ is PF_2 if and only if $g(z) > 0$ $\forall z \in \Re$ and $g(z)$ is log–concave on $\Re$.*

PROOF: See Appendix 1. ∎

The equivalence result of Lemma 1 already provides us with useful relationships summarized in the following proposition.

PROPOSITION 2: *The following conditions are equivalent:*
(a) $F_i(\theta_i)$ is IHR *in θ_i on $\{\theta_i \in \Theta_i : F_i(\theta_i) < 1\}$,*
(b) The survival function $\overline{F}_i(\theta_i) = 1 - F_i(\theta_i)$ is PF_2,
(c) The survival function $\overline{F}_i(\theta_i)$ is log–concave.

PROOF: Equivalence of (b) and (c) is a direct consequence of Lemma 1. From Definition 1 it follows that:

$$r_i(\theta_i) = \frac{-\overline{F}'_i(\theta_i)}{\overline{F}_i(\theta_i)} \qquad \text{on} \quad \{\theta_i \in \Theta_i : F_i(\theta_i) < 1\}, \tag{49}$$

which has to be increasing on Θ_i because by Lemma 1 the quotient $\overline{F}'_i(\theta_i)/\overline{F}_i(\theta_i)$ is decreasing on Θ_i as $\overline{F}_i(\theta_i)$ is log–concave. Thus (a) and (c) are equivalent. ∎

Therefore, if I can identify that the survival function of θ_i is log–concave, Proposition 2 ensures that its distribution $F_i(\theta_i)$ is IHR. This result suffices to ensure the existence of a separating equilibrium with quantity discounts in an optimal nonlinear pricing problem (provided that appropriate demand conditions summarized in Assumption 1 and Assumption 3 hold). The following lemma proves that IHR is preserved under convolution so that both, the *ex–ante* and the *ex–post* nonlinear pricing problems are consistent and well defined.

LEMMA 2: *Let $F_1(\theta_1)$ and $f_2(\theta_2)$ be PF_2, then the convolution* (46) *is also PF_2.*

---

[17] Total positivity can be understood as a generalization of convexity of functions. Many density functions, *e.g.*, all those of the exponential family, are totally positive, and are characterized by nice smoothness, sign regularity, and differentiability properties. Pólya frequency functions are single–peaked and have bounded rate of decay, what makes them particularly useful in reliability theory.

PROOF: Straightforward application of the *Basic Composition Formula*.[18] See Appendix 1. ■

So far, I have shown that the convolution of log–concave functions is also log–concave, and that if the involved function is the survival function, the convolution distribution is IHR. However, this result is of limited application because in most cases neither the hazard rate or the survival function have a simple closed form expression, and thus checking for log–concavity of survival functions or whether $r_i'(\theta_i) > 0$ is not generally feasible.[19] This is not the case with probability density functions since most of them have closed form expressions and log–concavity can be verified directly. The following result shows that the integral of log–concave functions is also log–concave, and therefore, a sufficient condition to ensure that $F_i(\theta_i)$ is IHR is to require log–concavity on the corresponding probability density function, $f_i(\theta_i)$.

PROPOSITION 3: *If $f_i(\theta_i)$ is PF_2, then the corresponding cumulative distribution function $F_i(\theta_i)$ and survival function $\overline{F}_i(\theta_i) = 1 - F_i(\theta_i)$ are PF_2.*

PROOF: See Appendix 1. ■

This result is very useful because it links the log–concavity of the probability density function $f_i(\theta_i)$ to the log–concavity of its distribution and survival functions. Then, as Corollary 1 shows, for the problem studied in this paper, it suffices to assume that the appropriate probability density function is log–concave to ensure that the corresponding nonlinear pricing problem is well behaved.[20] The following corollaries summarize the operative results that help characterizing the appropriate distribution of the asymmetric information parameters.

COROLLARY 1: *If the probability density function $f_i(\theta_i)$ is at least twice continuously differentiable and log–concave, the following properties are all equivalent:*
(a) $f_i(\theta_i)$ is PF_2,
(b) $F_i(\theta_i)$ is log–concave,
(c) $\overline{F}_i(\theta_i) = 1 - F_i(\theta_i)$ is log–concave,
(d) $F_i(\theta_i)$ is IHR *in θ_i on $\{\theta_i \in \Theta : F_i(\theta_i) < 1\}$.*

---

[18] The *Basic Composition Formula* is the continuous version of the Binet–Cauchy formula that expresses any minor of order k of the product of two rectangular matrices as the product of all possible minors of factors of order k. See Gantmacher (1959, §1.2.4–1.2.5) and Karlin (1968, §1.2).

[19] Some exceptions to this rule are the uniform, exponential, Weibull, and modified extreme value distribution, for all of which there exist closed form expressions for the hazard rate $r_i(\theta_i)$ in terms of the indexing parameters of each distribution. See Johnson, Kotz, and Balakrishnan (1995).

[20] The first part of Proposition 3 relating log–concavity of $f_i(\theta_i)$ and $F_i(\theta_i)$ is originally due to Prékova (1971) in a much more abstract setting not restricted to reliability operations. It is also possible to ensure that the optimal nonlinear pricing solution is well behaved by adopting the weaker assumption that the involved cumulative distribution function $F_i(\theta_i)$ is log–concave instead of $f_i(\theta_i)$ [Karlin (1968, §3.5)]. This weaker assumption is however of limited application because while shifting the log–concavity assumption from the density to the distribution function excludes few distributions from being considered [Barlow, Marshall, and Proschan (1963)], it makes operationally more difficult to identify monotonicity properties as explained before.

PROOF: Part (a) is a direct consequence of Lemma 1. Parts (b) and (c) follow from Lemma 1 and Proposition 3. Finally part (d) combines the results of Lemma 1, Proposition 2, and Proposition 3. ■

COROLLARY 2: *The cumulative convolution distribution function, $F(\theta)$, of two log–concave probability distribution functions, $f_i(\theta_i)$, $i = 1, 2$, is* IHR.

PROOF: By Lemma 1, if the distribution function $f_i(\theta_i)$ is log–concave, it is also PF_2 whenever it is positive and twice continuously differentiable, which holds by Assumption 2. Lemma 2 establishes that the convolution:

$$f(\theta) = \int_{\Theta_2} f_1(\theta - \theta_2) f_2(\theta_2) d\theta_2, \tag{50}$$

is PF_2. Then by parts (a) and (d) of Corollary 1, $F(\theta)$ is IHR. ■

Provided that the probability density functions of θ_1 and θ_2 are log–concave, Corollary 1 and Corollary 2 ensure that the cumulative convolution distribution function of the *ex–post* type is IHR. Given Assumption 3 on the distributions of θ_1 and θ_2, the properties of the distribution of θ required by Assumption 2 in order to characterize the optimal standard nonlinear pricing schedule in Section 2 are ensured to hold and are no longer necessary to be considered as an assumption, but rather as a direct consequence of Assumption 3 and Definition 3. I state here this sufficient condition as an assumption.

ASSUMPTION 5: The probability density functions $f_i(\theta_i) \geq 0$, $i = 1, 2$, are log–concave on $\Theta_1 \subseteq \Re$ and $\Theta_2 \subseteq \Re$ respectively.

Thus, both the *ex–ante* and *ex–post* pricing problems are consistent, well defined, and have a separating equilibrium involving quantity discounts. Observe that this result has been achieved with minimum additional assumptions relative to the standard nonlinear pricing problem. In the standard case, it is necessary to assume a common knowledge IHR distribution for the asymmetric information parameter that is unknown to the monopolist. Since for the optional tariffs case there are two components of the type, I have already shown that it suffices to assume that both common knowledge distributions are also IHR.

4.2 Implications

The definition of the *ex–post* type given in equation (15) together with the regularity conditions of the distributions of the type components discussed so far ensure that θ second order stochastically dominate θ_i, *i.e.*, $\theta_i \leq^{st} \theta$. This is a direct testable implication of the type–varying model. However, as we will see in Section 6.1, second order stochastic dominance does not generally suffice to compare the relative expected efficiency of *ex–ante* and *ex–post* nonlinear tariffs as higher or lower markups are inversely related to the magnitude of the hazard rate of the involved type distribution. Second order stochastic dominance allows non–uniform orderings of the markups of the *ex–ante* tariff relative to those of the *ex–post* tariff, depending on particular consumption ranges.

A sufficient condition to compare the optimal solutions of the *ex–ante* and *ex–post* nonlinear pricing mechanisms is to find and/or require a particular hazard rate ordering of the involved distributions. Since optimal nonlinear solutions critically depend on the value of the hazard rate of the corresponding distribution I have to establish how large is the hazard rate of the convolution distribution $F(\theta)$ relative to those of the components of the *ex–post* type. Proposition 4 shows that for the present type–varying model, θ dominates in hazard rate to θ_i if the support of the distributions is restricted to $\Re_+$.

PROPOSITION 4: *Let $F_i(\theta_i)$ be* IHR, *i.e., $r_i'(\theta_i) > 0$ in θ_i on $\{\theta_i > 0 : F_i(\theta_i) < 1\}$, for $i = 1, 2$. Let $F(\theta)$ denote the cumulative convolution distribution of $\theta = \theta_1 + \theta_2$, with hazard rate $r(\theta)$. Then $r(\theta) \leq \min\{r_1(\theta), r_2(\theta)\}$ on $\{\theta > 0 : F(\theta) < 1, F_i(\theta) < 1; i = 1, 2\}$.*

PROOF: See Appendix 1. ∎

The result of Proposition 4 implies that the distribution of θ always puts more weight on higher values than the distribution of θ_1. Therefore given some value $\hat{\theta}$, the probability that $\theta > \hat{\theta}$ always exceeds the probability that $\theta_1 > \hat{\theta}$. This intuitive result is formalized in the following corollary.

COROLLARY 3: *If $r(\theta) \leq r_i(\theta)$ on $\{\theta > 0 : F(\theta) < 1, F_i(\theta) < 1; i = 1, 2\}$, then θ first order stochastically dominates θ_i.*

PROOF: Since $r(\theta) = -d \log[1 - F(\theta)]/d\theta$ it follows that $\forall\theta > 0$:

$$1 - F(\theta) = \exp\left[-\int_0^\theta r(z)dz\right] \geq \exp\left[-\int_0^\theta r_i(z)dz\right] = 1 - F_i(\theta), \tag{51}$$

and therefore $F(\theta) \leq F_i(\theta)$ $\forall\theta > 0$, which is the definition of first order stochastic dominance, of θ over θ_i, *i.e.*, $\theta_i \leq^{st} \theta$ [Marshall and Olkin (1979, §17.A)]. ∎

Following Laffont and Tirole's interpretation (1993, §1.4–1.5), Proposition 4 means that the distribution of θ is more favorable than the distribution of θ_1. Corollary 3 shows that this result could be obtained within the type–varying framework because of the existence of an independent, but systematically positive, type shock ensures that the actual purchase (or valuation) is always higher in stochastic sense than the expected purchase (or valuation). Similarly, Maskin and Riley (1984, §4) already considered the effect of changes in the distribution of consumer types on the shape of the nonlinear tariffs. As I show in Section 6.1, a nonlinear schedule based on $F(\theta)$ generally involves higher markups than the nonlinear tariff based on $F_1(\theta_1)$ for all consumption levels, which is a direct consequence of the hazard rate dominance of θ over θ_1 in a model with type varying consumers if the support of the distributions are restricted to $\Re_+$.

The testable implications of Proposition 4 and Corollary 3 are, strictly speaking, limited to situations where $\theta, \theta_i \in \Re_+$, which exclude the empirical application of this paper. There is no reason to expect that consumers always underestimate their future local telephone usage, and thus the type shock θ_2 is not restricted to take only positive values. However, a strong empirical evidence of first stochastic dominance will be consistent with the underlying hazard rate dominance of θ over θ_1. Thus, in such a case, an *ex–post* nonlinear tariff could lead to higher expected profits than an *ex–ante* nonlinear tariff because the markups of $\hat{T}(\theta)$ uniformly dominates those of $\tilde{T}(\theta_1, \theta_2)$ for every consumption level.

Some other pricing or agency problems could however define environments where the support of type components is constrained in a natural way. For instance, we could think of $\theta_1 \in \Re_+$ as general skills of workers before being hired (*e.g.*, acquired through education and/or working experience in other jobs). If hired, workers could develop some specific skills and abilities due to learning by doing, and therefore increase their productivity. It is not unreasonable within this framework to exclude the possibility of negative learning, and thus θ_2 could also be restricted to take only positive values. Consumption of electricity also provides a related example. While households consume according to their habits and location, *i.e.*, the base load $\theta_1 \in \Re_+$, changes in temperature (public information and common to all consumers) may induce additional seasonal demand: whenever it is too cold or too warm, consumers increase their demand for electricity by turning on the heating or the air conditioner. Thus, θ_2 could also be restricted to take only positive values, and the model will produce stronger empirical implications.[21]

4.3 General Distributions and Type Definitions

In this subsection I discuss whether the generality of my results depends on the assumption of independence of type components and/or on the additive definition of the *ex–post* type.

Equation (46) establishes that given the distribution of θ_1 and θ_2 it is always possible to identify the distribution of θ up to a linear transformation of the distributions of the type components. This identification issue rises the question of whether the results of the model are limited to a particular definition of the *ex–post* type in terms of the *ex–ante* type and the type shock. It might be useful to discuss briefly the implications of this functional form identification constraint because in principle we could be interested in other specifications for the distribution functions of the form $G_i(\theta, \theta_i)$ where contrary to what happens in (46), θ is not necessarily a translation parameter. I find two types of arguments to ignore this alternative.

First, Statistics does not provide us with robust results concerning composition distributions. Eaton and Perlman (1991) show that even considering Schur–concave distribution functions $G_i(\theta, \theta_i)$, composition distribution are only well behaved depending on the system of coordinates chosen (*e.g.*, logarithmic *vs.* linear type definiton).[22] Furthermore, although for some products and quotients of random variables composition distributions

[21] On state contingent tariffs of this nature see Panzar and Sibley (1978) and Spulber (1992).

[22] For simplicity, assume that $\theta, \theta_i \in \Re$. Then, a distribution function $G_i(\theta, \theta_i)$ is said to be Schur–concave if and only if (i) it is permutation symmetric, *i.e.*, $G_i(\theta, \theta_i) = G_i(\theta_i, \theta)$; and (ii) $(\theta - \theta_i)[\partial G_i(\theta, \theta_i)/\partial\theta - \partial G_i(\theta, \theta_i)/\partial\theta_i] \leq 0$. See Marshall and Olkin (1979, §1.C and §3.A).

are well defined, there are no general results relating the hazard rate properties of the composed distributions.[23]

Second, and more importantly, in Consumer Theory the distinction between convolution and composition distribution is mostly irrelevant because it does not rely on particular functional forms of the utility function. It is well known that any monotone transformation of a quasi–concave utility function represents the same set of preferences and provides with exactly the same comparative statics analysis. Similarly, the Economics of Information has dealt mostly with a single–dimensional parameter to index agents' preferences, and assumed that the distribution of this index is common knowledge in order to characterize (as a Perfect Bayesian–Nash Equilibrium) the principal's optimal mechanism in the presence of adverse selection and/or moral hazard. Within the framework of the present paper, all this implies that the functional form that relates the *ex–ante* type and the shock with the *ex–post* type is not independently identifiable from the assumed distributions and/or the utility function that represents consumers' preferences. Since the preference index that represents consumer's types is just a theoretical construction to describe situations of asymmetric information, we can always define a monotone transformation of the utility function that scale the index and its distribution appropriately to represent the same preferences, which ensures the generality of the results of the present model.[24] In the end, types shocks, regardless of whether they are linearly related to the *ex–ante* type or not, are only identifiable as non–price related shifts in consumer demand. Therefore, I can focus on the convolution case to analyze, without loss of generality, the implications of the existence of type shocks for the design of optimal pricing mechanisms.

Let me finally address the issue of the independence of θ_1 and θ_2. From an economic point of view it is interesting to consider cases where the *ex–ante* type and the shock are correlated. It is not difficult to envision situations where large consumers also make more or less mistakes than small consumers. In principle we should expect that households with many members are more difficult to monitor and that intensive users of telephone services also experience more significant consumption variations over time than small consumers with more stable consumption patterns. Although the present paper does not suggest any theory to explain the positive or negative correlation between the *ex–ante* type and the

[23] See Marshall and Olkin (1979, §3.J) and also Springer (1979) for a general treatment of products, quotients, and algebraic functions of random variables using the Mellin integral transform. Miravete (1997) develops an empirical application based on the product of two beta distributed variables.

[24] For instance, assume that the utility function of consumers is multiplicatively separable in the *ex–post* type, θ, and consumption. Assume also that the *ex–post* type is the product of the *ex–ante* type, θ_1, and the shock, θ_2, each of which is beta distributed with appropriate parameters so that the composition distribution is uniform (a particular case of beta distribution). Observe that the same set of preferences can be represented by a logarithmic transformation of that utility function. But in that case, the asymmetric information parameter and consumption are additively separable in the new utility function and the *ex–post* type –now distributed according to a $\chi^2(2)$– becomes a linear combination of the *ex–ante* type and the shock, now both distributed as exponential generalized beta of the second kind. For the model to work, these latter distributions should also be log–concave. See Hogg and Craig (1995, §4.3), Kotlarski (1962), and McDonald and Xu (1995).

shock, it might be interesting to analyze how would those cases be related to the results presented here.

For example, consider the reference case where θ_1 and θ_2 are independent and $\theta \leq^{st} \theta_1$, or equivalently, $F(\theta) \leq F_1(\theta)\ \forall\theta \in \Theta \subseteq \Re$. Assume now that the *ex–ante* type and the shock are negatively correlated, and denote by $F^\star(\cdot)$ the cumulative distribution of $\theta = \theta_1 + \theta_2$ under negative correlation. The distribution of θ is now less dispersed, with less mass of probability at the tails of the distribution than if θ_1 and θ_2 were independent. In some sense the monopolist is now "less uncertain" about the value that consumer types may take, because there is a larger mass of probability around the mean of θ. Thus, for low values of θ (below the mean), the probability of finding values of the type above a given θ is higher under negative correlation than under independence. This is just because the probability distribution function of $\theta_1 + \theta_2$ is more concentrated around the mean. Thus, as the survival function is higher, the hazard rate function is lower under negative correlation than under independence for low values of θ. Just the contrary holds for high values of θ (above the mean). Under negative correlation there is less mass of probability left above the mean than under independence of type components. Thus, for high types, the hazard rate of the distribution with negative correlation will exceed that of the distribution of independent type components. If $r^\star(\theta) \leq r(\theta)$ only for low values of θ, then for large customers *ex–post* nonlinear pricing markups will be lower under negative correlation of type components than under the assumption of independence as markups and hazard rate of the distribution of θ are inversely related.[25] Consumer types are more concentrated around the mean under negative correlation than under independence, and thus it is necessary to introduce important distortions to distinguish among low consumers and preserve the IC property of the mechanism.[26] Thus, the results of this paper in general, and of this section in particular need to be qualified for particular cases where type components are allowed to be correlated.

The results of this section could still hold for a modified version of the problem. The modified problem requires a change of basis for the space of asymmetric information parameters, $\Theta_1 \times \Theta_2$. The new basis $\Theta_1' \times \Theta_2'$ should be defined through spectral decomposition of the covariance matrix of types (θ_1, θ_2), so that the covariance matrix of the new types (θ_1', θ_2') is diagonal. Technically, the *Basic Composition Formula* could then be applied and the results of this section would remain valid using the new types (θ_1', θ_2'). However, this new tractable problem has not obvious economic interpretation because

---

[25] See Section 6.1 below.

[26] These intuitive results are however difficult to prove except, maybe, for particular distributions. The reason is that under correlation, the probability density function of $\theta = \theta_1 + \theta_2$ cannot be factorized as the product of the components' probability density functions, and thus $F(\theta)$ is no longer the convolution distribution of θ_1 and θ_2. The distribution of the sum of non–independent random variables is generally difficult to obtain in closed form. Most cases have to be approximated by quadrature methods [Polyanin and Manzhirov (1998, §8.7)]. One of the cases where $F^\star(\theta)$ can be written explicitly is that of $\theta = \theta_1 + \theta_2$ where $(\theta_1, \theta_2) \sim BVN[\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho]$. In this case, $\theta \sim N[\mu = \mu_1 + \mu_2, \sigma^2 = \sigma_1^2 + 2\rho\sigma_1\sigma_2 + \sigma_2^2]$ [Springer (1979, §3.4)]. To illustrate the argument of this paragraph, I computed the hazard rate functions of $\theta_1 + \theta_2$ under independence, $r(\theta)$, and under perfect negative correlation ($\rho = -1$), $r^\star(\theta)$. For the case where $\mu_1 = 0$, $\sigma_1^2 = 1$, $\mu_2 = 1$, $\sigma_2^2 = 0.5$, I found that $r^\star(\theta) > r(\theta)\ \forall\theta > 0.12$.

while θ_1 and θ_2 clearly isolate the permanent and stochastic components of consumers' tastes, θ_1' and θ_2' are defined as linear transformations of θ_1 and θ_2 and therefore both enter the definition of the new type components in different proportions. Contrary to the identification issues treated before, dealing with the correlation of types actually affects the nature of the economic problem studied, and the distinction between permanent and stochastic aspects of the types becomes meaningless when we analyze the problem through principal component analysis.

5 Empirical Evidence

Results of Section 4 provide us with direct testable implications of the "taste–varying" approach. The goal of this section is to test whether the distribution of *ex–post* types second order stochastically dominates the distribution of *ex–ante* types. Contrary to many applied works, the source of asymmetric information in the application studied here is not identified through the specification of some distribution of unobserved characteristics, but rather using direct observations of consumers' taste parameters. The empirical analysis exploits the information available from the 1986 Kentucky Local Telephone Tariff Experiment.

In November of 1984, the Kentucky Public Service Commission (KPSC) established Administrative Case #285 to study the economic feasibility of providing local measured service telephone rates. Directly linked to Case #285, SCB carried out an extensive tariff experiment in the second half of 1986 in two cities of Kentucky to provide the commission with evidence in favor of introducing the optional local measured service. Prior to this tariff experiment, in spring, when all households in Kentucky were on mandatory flat rates, SCB collected demographic and economic information for about 5,000 households in the local exchanges of Bowling Green and Louisville. The regulated monopolist also collected monthly information on usage (number and duration of calls classified by time of the day, day of the week, and distance), and payments during two periods of three months in spring and the fall of that year.

It is remarkable that in addition to demographic and economic variables, SCB also collected information on telephone customers' usage expectations. SCB explicitly requested customers' own estimates of their weekly average number of calls. These individual estimates are particularly useful because local calls were never priced before and consumers were not aware of the tariff experiment that was going to be held in the second half of the year. Thus, neither marginal tariffs or strategic considerations influence these estimates of their own satiation levles. This information, available for most households of the sample can be compared with the actual number of weekly phone calls for every month in the study. A direct test of the suggested taste–varying model will be constructed in the next subsection using the comparison between the expected and the actual weekly number of phone calls. Appendix 2 describes in detail all variables used in this study and Table 1 presents basic descriptive statistics of the sample stratified by location.

These two cities have quite different demographic structures. Residents in Bowling Green make a significantly higher income and households are larger, including the proportion of teenagers. Households with married couples and college graduates are also more common in Bowling Green than in Louisville. In this latter city, on the contrary, it is more common to find retired people, those who receive some kind of social benefits to support their income, and a smaller percentage of households that have moved in the last five years. Racial composition of these cities is also different. Only 6% of the population in Bowling Green, but about 12% of the population in Louisville, is black.

There is also a significant difference between usage and expected usage of local telephone service across these two local exchanges. While consumption (measured as weekly number of calls) is higher in Louisville than in Bowling Green, the expected consumption is much more accurate in this last exchange. On average, Bowling Green residents underestimate telephone usage by 2%, Louisville residents underestimate their usage by 29%. The difference in magnitude of the bias (type shock of the model) is remarkable. Perhaps it could be explained by positive network effects of the size of the local exchanges [Taylor (1994, §9)]. While Bowling Green barely reached 50,000 inhabitants by the end of the 1980's Louisville had a population that exceeded 250,000.

5.1 Are Data Consistent with the Type–Varying Model?

A common problem in estimating demand when consumers face nonlinear budgets is that the choices of consumption and the marginal tariff are simultaneous and therefore the relevant price is endogenous [*e.g.*, MacKie–Mason and Lawson (1993, §3.2)]. Regarding this point, observe that comparing the expected weekly number of calls with the actual number of calls during the spring months is qualitatively different from comparing those expectations with the actual number of calls during the fall months. In the second case the number of phone calls is a function of the tariff chosen (in Louisville) and the marginal charge per call, which varies with the time of the day and distance of the outgoing call (both in Bowling Green and Louisville's measured service option), as well as of customers' accumulated monthly usage of telephone services. However, this is not the case during the spring months because all local telephone customers were placed under a mandatory flat rate regime. Price was a relevant economic variable for the decision to subscribe the telephone service, but once it had been subscribed any additional call involves a zero marginal charge, and consequently local telephone customers should consume at their satiation levels.[27]

Focusing on the spring months, the present data set provides us with an uncommonly available direct indicator for θ_1, the expected number of weekly calls, and also for θ, the actual number of weekly calls. Provided that the mean of θ_2 is finite, *i.e.*, $E[\theta_2] = \mu_2$,

[27] According to the 1990 U.S. Census of Population and Housing, 89.33% of the households in Bowling Green and 92.07% of those in Louisville were subscribed to local telephone service. The assumption that the whole market is served by the monopolist is therefore not unrealistic.

and given that θ is additively separable in θ_1 and θ_2, –*i.e.*, equation (15)–, $F(\cdot)$ will be a stochastic spread of $F_1(\cdot)$, and therefore θ_1 should second order stochastically dominate θ. The stochastic spread will be mean preserving, mean increasing, or mean decreasing depending on whether μ_2 is zero, positive, or negative. [28]

The first column of Table 2 shows the average usage expectation bias, μ_2, for residents in Bowling Green and Louisville respectively. The average bias is positive for customers of these two local exchanges, but it is about seventeen times larger in Louisville than in Bowling Green. A more detailed analysis by demographic strata shows further differences between residents of these two exchanges. While in Louisville the bias is always positive and large, independently of the demographic characteristic considered, in Bowling Green it is more balanced and in several occasions the average bias takes negative values. In both cities consumers tend to underestimate their future usage, but in Louisville they do this by more than an order of magnitude. The smaller average bias in Louisville (single and male household) is still more than seven times larger than the average bias in Bowling Green. Figures 1.5–1.6 show the empirical density function of the type shock. Although these expectation bias are quite disperse (the standard deviation for Bowling Green is about 37 calls per week while in Louisville reaches almost 40 calls), small mistakes around the mean are the most frequent event. Thus, the excess of kurtosis is 51.31 for Bowling Green and 42.98 in Louisville. But since customers in Louisville have a considerably more biased expectation (on average 10.7 calls *vs.* 0.6 calls in Bowling Green), their empirical distribution of shocks is positively skewed, 1.08, while in Bowling Green is negatively skewed, -4.09. The "PAT" column presents further evidence in favor of the type varying model by computing Pearson's analog goodness of fit test for the equality of $F(\cdot)$ and $F_1(\cdot)$. That hypothesis is always strongly rejected and therefore we can conclude that the distribution of θ_2 is not degenerate and that the suggested type varying model is an accurate representation of consumers preferences.

There is significant heterogeneity by strata for local telephone usage expectation bias. In both cities there is evidence (stronger in Louisville) in favor of a mean increasing spread of the distribution of θ relative to that of θ_1. However, a systematic ordering of the means of θ and θ_1 (through a positive μ_2) is not sufficient to ensure the stochastic dominance of θ over θ_1, since the whole distribution matters [Mas–Colell, Whinston, and Green (1995)]. Figures 1.1–1.2 present the empirical frequency distributions of actual and expected weekly number of local calls for the spring months of the experiment in the local exchanges of Bowling Green and Louisville respectively. It is evident that the distribution of expected weekly calls is characterized by the accumulation of frequencies on a few "focal points" of the usage range. While the concentration of probability around these focal points could partially explain that telephone customers underestimate their future consumption, there is no reason why these "focal points" could not be shifted upwards some few units (especially in the case of Louisville) to provide a less biased estimator. More informative

[28] Second order stochastic dominance (SOSD) will occur even if the shock has a mean different from zero. See Laffont (1993, §2.5).

is however the empirical cumulative distribution functions shown in Figures 1.3–1.4, which clearly indicates that in both cities telephone customers tend to underestimate their future local telephone usage over the whole calling range, which leads to the relative ordering of the averages of θ and θ_1 discussed in Table 2.

The testable implication of the suggested type–varying model is that θ, the number of actual weekly calls stochastically second order dominates θ_1, the expected number of weekly calls. Figure 1.4 appears to indicate that θ first order stochastically dominates θ_1 in Louisville, although Figure 1.3 fails to prove the same for Bowling Green. While FOSD implies SOSD and therefore supports the suggested type–varying model, FOSD is also much more restrictive than SOSD because it implies that consumers systematically underestimate their future consumption, not only independently of their demographic characteristics, but also independently of the magnitude of their local telephone usage. Thus, strong FOSD of θ over θ_1 will be consistent with a model where θ dominates in hazard rate to θ_1. The distribution of the actual number of calls will be more favorable than the distribution of expected number of calls, and therefore the optimal *ex–post* type based nonlinear price schedule would imply uniformly higher markups than the *ex–ante* type based nonlinear price schedule for all possible local telephone usage levels.

In order to test the hypotheses of first and second stochastic dominance, I computed Anderson's (1996) nonparametric test of stochastic dominance. The test, whose construction is briefly summarized in Appendix 3, is based on comparing weighted differences of frequency functions of two variables within given mutually exclusive fractiles. The null hypothesis of FOSD of θ over θ_i, *i.e.*, $\theta_i \leq^{st} \theta$, requires that:

$$F(\theta) \leq F_i(\theta), \qquad F(\theta') \neq F_i(\theta') \quad \text{for some } \theta',\ \forall \theta \in \Theta. \tag{52}$$

Similarly, SOSD of θ over θ_i requires that:

$$\int_{\underline{\theta}}^{\theta} [F(z) - F_i(z)]\, dz \leq 0, \qquad F(\theta') \neq F_i(\theta') \quad \text{for some } \theta',\ \forall \theta \in \Theta. \tag{53}$$

Test results presented in Tables 3 and 4 compare these two inequalities for a given range of phone calls (20 fractiles). Stochastic dominance of any given order requires that each inequality holds for every given call category, *i.e.*, not a single test of each row can be significantly positive. Tables 4a–4b provide with strong evidence in favor of the suggested type–varying model, as SOSD of θ over θ_1 is never rejected in neither of the two cities and for any of the demographic strata.[29]

Absolute SOSD test values are systematically higher in Louisville than in Bowling Green. This is partly due to the different behavior of $F(\cdot)$ relative to $F_1(\cdot)$ in these two

[29] I furthermore checked that SOSD was never rejected for neither of the two cities in any single month, using 10 and 15 fractiles.

cities. As Tables 3a–3b report, FOSD of θ over θ_1 is rejected in Bowling Green (for households with high income, young, and single and male head, as well as for the high end of the calling range), but never in Louisville. This FOSD result explains why the type shock has always a positive mean in Louisville, and its rejection in Bowling Green is consistent with the negative average bias found for some demographic strata.

6 Welfare Analysis

In this section I study the possible welfare effects due to the introduction of optional nonlinear tariffs. First, I explore the possibilities of obtaining general results from a theoretical perspective. Second, I compute estimates of the welfare components using the empirical distributions of θ and θ_1 of the Louisville and Bowling Green samples.

6.1 Freedom of Choice vs. Mandatory Pricing

After the breakup of AT&T optional pricing became a common practice in Telecommunications and also in other industries. Two questions at least arise. First, why do firms find profitable to make optional tariffs available to their customers? And second, from the Regulator's perspective, should optional pricing be allowed in regulated industries? The analysis of this section intends to shed some light on this complex issue.

Comparing the solution of the *ex–post* and *ex–ante* problem it is straightforward to realize that there are two elements that characterize the solutions of the optimal marginal tariffs: the hazard rate of the appropriate distribution for each problem, and how the demand conditions enter the solution, *i.e.*, for the *ex–ante* problem, the effects of the type shock θ_2 are integrated out while for the *ex–post* tariff this is not the case.

I will start by analyzing the role of asymmetric information in the solution of the optimal *ex–post* nonlinear pricing problem. Suppose that consumer type θ could be distributed with respect to either $F(\theta)$ or $G(\theta)$. Which of these two distributions is more informative for the monopolist? In the previous sections, I have repeatedly made use of an inverse relationship between the hazard rate of the distribution and the mark–up that the monopolist can charge to his customers. The following Proposition shows the important role that the hazard rate of the distribution plays in the discussion of profitability of different pricing schemes.[30]

Proposition 5: *Let $F(\theta)$ and $G(\theta)$ be* IHR, *i.e., such that $r'_F(\theta) > 0$ in θ on $\{\theta > 0 : F(\theta) < 1\}$, and $r'_G(\theta) > 0$ in θ on $\{\theta > 0 : G(\theta) < 1\}$. Assume also that $r_F(\theta) \leq r_G(\theta)$, $\forall\theta$. Then, the price mark–up and the marginal tariff will be uniformly higher under the $F(\theta)$ distribution than under the $G(\theta)$ distribution.*

[30] Different versions of this result have been previously stated by Ivaldi and Martimort (1994), Laffont and Tirole (1993, §1.5), and Maskin and Riley (1984, §4).

PROOF: From solution (9) of the *ex–post* pricing problem, the price mark–up is:

$$\frac{\hat{p}(\theta)-c}{\hat{p}(\theta)} = \frac{-r^{-1}(\theta)\left[\dfrac{v_{p\theta}(p(\theta),\theta)}{v_{pp}(p(\theta),\theta)}\right]}{c-r^{-1}(\theta)\left[\dfrac{v_{p\theta}(p(\theta),\theta)}{v_{pp}(p(\theta),\theta)}\right]}. \tag{54}$$

Differentiating this expression with respect to $r(\theta)$ we obtain:

$$\frac{\partial}{\partial r(\theta)}\left(\frac{\hat{p}(\theta)-c}{\hat{p}(\theta)}\right) = \frac{c\left[\dfrac{v_{p\theta}(p(\theta),\theta)}{v_{pp}(p(\theta),\theta)}\right]}{r^2(\theta)\left[c-r^{-1}(\theta)\left[\dfrac{v_{p\theta}(p(\theta),\theta)}{v_{pp}(p(\theta),\theta)}\right]\right]^2} < 0. \tag{55}$$

Thus, evaluating these expresions at $r_F(\theta)$ and $r_G(\theta)$ respectively, it follows that the price mark–up will be higher under $F(\theta)$ than under $G(\theta)$. ▪

The monopolist will charge a higer price mark–up under $F(\cdot)$ than under $G(\cdot)$ because $F(\cdot)$ dominates in hazard rate to $G(\cdot)$. This is because distribution $F(\cdot)$ puts more weight on consumers of high type (closer to $\overline{\theta}$).[31] The fact that distribution $F(\cdot)$ is more favourable than $G(\cdot)$ implies that the optimal pricing has to create stronger incentives for inframarginal consumers to self–select according to their true type. Proposition 6 shows that maintenance of the IC constraint under more favourable distributions requires higer price distortions for inframarginal consumer types. Thus, reducing consumers' expected informational rents, the monopolist is able to screen among the many particular type values that a smaller proportion of his customers can reveal through their tariff choice or usage decisions.

But in the problem analyzed in this paper, substituting one distribution for a different one is just one factor, but not the only one, driving the welfare results. As it is obvious from the comparison of the marginal tariffs of the optimal *ex–post* and *ex–ante* nonlinear tariffs –equations (9) and (26)–, they will only differ on the distribution used if the ratio $v_{p\theta}/v_{pp}$ is independent of θ_2.[32] Comparing information structures that lead to the hazard rate ordering provides with a unique example in which different nonlinear tariffs can be sorted. But more frequently, comparison among informational structures will not lead to situations in which one distribution is more favorable than the other over the whole support of the distribution of types. This is the case of the present case study. Figure 2 compares the hazard rate of the distributions of expected and actual number of calls in

[31] Strictly speaking, this is true when the support of the distribution is restricted to $\Re_+$, because, as it has been shown in Corollary 3, then $F(\cdot)$ would first order stochastically dominate $G(\cdot)$. Otherwise, hazard rate dominance is a regularity condition that has to be imposed in order obtain that markups under $F(\cdot)$ uniformly dominate those under $G(\cdot)$ even when $F(\cdot)$ FOSD $G(\cdot)$.

[32] This condition could be fulfilled by some multiplicative specifications of θ. In particular utility functions where the *ex–post* type enters as a scale factor only, *e.g.*, indirect utility functions that are multiplicatively separable in types and prices.

the two local exchanges. In both cases, there is an alternating dominance in hazard rate of $F(\cdot)$ over $F_1(\cdot)$ and *vice versa*. But in Louisville, consistent with the FOSD result, $F_1(\cdot)$ dominates in hazard rate to $F(\cdot)$ only for very small ranges: from approximately 52 to 59 and from 85 to 105 calls per week. Thus, without strict hazard rate dominance markups can be higher under one tariff only for a given range of consumption. This result is even more likely to happen when we also consider that demand conditions enter differently under the optimal *ex–ante* and *ex–post* nonlinear tariffs. The question that remains to be answered is whether, regardless of all these issues, something can be said about the desirability of *ex–ante vs. ex–post* tariffs for different agents.

The results for the monopolist are conclusive. Optimal tariff functions $T(\theta)$ are necessarily increasing, $T'(\theta) = p(\theta) > 0$. Furthermore, if the problem is well behaved, tariff functions will be concave (quantity discounts), $T''(\theta) = p'(\theta) < 0$. Therefore, the monopolist generally expects an increase in profits by introducing optional pricing. The result is a direct consequence of the classical conditions of Hadar and Rusell (1969) to order outcomes under uncertainty.

PROPOSITION 6: *Expected profits are higher under ex–ante pricing if any of the folowing conditions hold:*
(i) $T'(\theta) > 0$ *and* $F(\cdot)$ *FOSD* $F_1(\cdot)$,
(ii) $T'(\theta) > 0$, $T''(\cdot) < 0$, *and* $F(\cdot)$ *SOSD* $F_1(\cdot)$.

PROOF: Under circumstances of part *(i)*, the difference of expected profits between *ex–post* and *ex–ante* tariffs is (integrating by parts):

$$\int\limits_{\Theta} T(x)[F(x) - F_1(x)]dx = -\int\limits_{\Theta} T'(x)[F(x) - F_1(x)]dx \geq 0, \tag{56}$$

while for part *(ii)* the result is obtained integrating (56) by parts again:

$$\int\limits_{\Theta} T''(x) \int\limits_{\Theta} [F(y) - F_1(y)]dydx - T'(x) \int\limits_{\Theta} [F(y) - F_1(y)]dy \Big|_{x=\underline{\theta}}^{x=\overline{\theta}} \geq 0, \tag{57}$$

which completes the proof. ■

Therefore, more favorable distributions (FOSD) increase expected profits even in situation in which the pricing problem does not fulfill all required conditions to discriminate among consumers by means of quantity discounts. But if these quantity discounts are optimal, then less restrictive stochastic orderings (SOSD) also lead to the same conclusion. The commonly observed practice of using optional nonlinear tariffs is therefore profit maximizing under very general conditions, which should suffice to explain its widespread use.

Unfortunately, I cannot affirm the same about consumers. Assumption 1 only requires that the indirect utility function be increasing in θ. But the effect on the net rent $v(p(\theta), \theta) - T(\theta)$ will depend on many factors. If $v(\cdot)$ is more increasing than $T(\cdot)$, then part *(i)* of Proposition 6 could be applied, and consumers will prefer optional pricing to mandatory *ex–post* pricing. However, it is possible to observe that type shocks are so biased that $F(\cdot)$ first order stochastically dominates $F_1(\cdot)$ and consumers still prefer the mandatory measured service. There would not be any assumption violation. Preferences just fail to be increasing enough in θ. This actually happens for the case of Louisville in the empirical analysis of the next section. A similar analysis could be made for the case of SOSD in order to apply part *(ii)* of Proposition 6. In addition to $v(\cdot)$ being more increasing than $T(\cdot)$, it would now require that $v(\cdot)$ is more concave than $T(\cdot)$. Thus, even more restrictive preferences are necessary to obtain a definite ordering of pricing strategies under increasingly less restrictive stochastic environments.

Obviously, this difficulty in ordering pricing schemes according to the expected consumer surplus is translated to the regulator's welfare measure. The result is more unclear the more weight is given to consumers in the regulator's objective function. But still something can be said, at least in limiting cases. Consider the following family of convoluted distributions defined according to (46) such that:

$$F(\theta, \theta_1, \theta_2, \tau) \qquad ; \qquad \tau = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \qquad ; \qquad F_\tau(\theta, \theta_1, \theta_2, \tau) > 0. \tag{58}$$

Thus, for instance, if $\sigma_2^2 = 0$ and $\tau = 0$ then $F(\theta) = F_1(\theta_1)$. Obviously, if the variance of the shock is zero, all consumer differences are captured by the distribution of the *ex–ante type*. Therefore, there is no real distinction between tariff choice and usage decision. All consumers and also the monopolist would be indifferent between an *ex–ante* and *ex–post* tariffs. A model like this is actually equivalent to one where consumers are able to commit *ex–ante* to their future consumption.

Without being too precise I should conclude that the smaller is the variance of the shock relative to the variance of the ex–ante type, the more likely is that *ex–ante* tariffs are welfare increasing and *vice versa*. To confirm this intuition, the other extreme case should also be analyzed. If $\sigma_1^2 = 0$ and $\tau = 1$, then $F(\theta) = F_2(\theta_2)$. In this case, consumers only differ *ex–post*. Since consumers are all alike *ex–ante*, the optimal *ex–ante* tariff will be a two–part tariff. But that is not the welfare enhancing (neither profit maximizing) tariff because consumers will be considerably more diverse *ex–post* than *ex–ante*, and both welfare and profits are increasing in the number of self–selecting tariffs [Faulhaber and Panzar (1977, §4); Wilson (1993, §8)], which implies that the optimal strategy should be an *ex–post* based fully nonlinear tariff.

This discussion just indicates whether a particular type of tariff is more likely to be welfare enhancing than another, based on the welfare increasing effect of more numerous self–selecting options when consumers are diverse. However, the discussion falls short of determining the threshold levels of the ratio σ_1^2/σ_2^2 that make *ex–ante* pricing dominate *ex–post* tariffs or *vice versa*. Determinants of this threshold are application specific, and will critically depend on the specification of the utility function and the distributions used.

Finally, observe that more general convolution families than those in (58) could be defined. For instance, the correlation between θ_1 and θ_2 could also be considered, although in practice we will face the problems discussed in Section 4.3.[33] But consideration of non–independent type components may also help characterizing whether *ex–ante* tariffs represent a welfare improvement relative to *ex–post* measured services. The idea is that there is some sort of relationship between hazard rate dominance and the dispersion of the distributions. The net effect depends on whether the nature of the shock increases or decreases the heterogeneity (asymmetric information) of consumers. For instance, we can return to the example discussed in Section 4.3 to illustrate the difference between optimal nonlinear tariffs in environments with and without correlation between type components. If type components are positively correlated, then consumer diversity will be more intense *ex–post*. This case has already been addressed before, when $\sigma_1^2 > \sigma_2^2$ (in particular the polar case $\sigma_2^2 = 0$). However, the negative correlation case add something substantial that previous cases could not address. In the case of negative correlation consumers are more heterogeneous *ex–ante* than *ex–post*. If shocks are such that consumers become so similar *ex–post* that $F(\theta)$ is close to degenerate, it is obvious that the welfare maximizing strategy consists in offering optional tariffs.

Thus, the stochastic nature of type shocks, whether they make the *ex–post* distribution more favorable (hazard rate effect) or whether consumers become more or less heterogeneous (variance effect) appear to drive welfare effects in particular applications. In general we can conclude that welfare effects will critically depend on the relative size of the variance of θ_1 and θ_2 and how they enter the definition of the hazard rate of the distributions of θ and θ_1 respectively. These relationships are complex enough to make precise predictions almost impossible. Perhaps, the exception is any family of distributions that can be ranked according to their hazard rates over the whole support $\Re$, regardless of the variances of the components of the type. Otherwise optimal *ex–ante* and *ex–post* tariffs will intersect each other, which therefore leads to ambiguous welfare results. Instead of solving the model for one of such cases that avoids ambiguity, the following section explores the magnitude of the welfare effects associated to different pricing strategies using the empirical distributions of θ, θ_1, and θ_2, which is a more interesting analysis because direct observations of types are rarely available.

6.2 Simulations

In this subsection I compute simple estimates of the welfare effects due to the introduction of optional tariffs for the two local exchanges of Kentucky using Monte–Carlo simulations.[34] For each city I evaluate the average expected consumer surplus, profits, and total welfare

[33] The evaluation carried in footnote 26 is an approximation to the effects of changes in the correlation parameter that could be added to the discussion here.

[34] Although optional tariffs were effectively introduced only in Louisville in the second half or 1986, there is no reason not to evaluate the potential effects of the introduction of optional tariffs in Bowling Green.

of screening local telephone customers through either a mandatory *ex–post* pricing, a continuum of optional two–part tariffs, or a continuum of fully nonlinear options. In order to evaluate the welfare effects, I first have to specify a particular demand function. I assume that the indirect utility function is:

$$V(p, A, \theta) = \frac{\theta}{\alpha} \exp[-\alpha p] - A \quad ; \quad \alpha > 0, \tag{59}$$

which leads to the following demand equation:

$$x(p, \theta) = \theta \exp[\alpha p]. \tag{60}$$

This specification has been used before in telecommunications demand analysis because it is bounded under the flat rate option [Hobson and Spady (1988); Kling and Van Der Ploeg (1990)]. If $p = 0$, consumers purchase their satiation level $x(p, \theta) = \theta$, which I identified in Section 5 as the actual number of weekly calls during the spring months when all customers were under mandatory flat rate service. Similarly, when $p = 0$, the expected usage equals $E_2[\theta] = \theta_1 + \mu_2$, which I identified as the expected number of weekly calls. The solutions of $\hat{T}(\theta)$, $\tilde{T}(\theta_1)$, and $\tilde{\tilde{T}}(\theta_1, \theta_2)$ for this particular demand function and general distribution, are shown in Appendix 3. Obviously, these equations could be solved explicitly for some given distributions. However, instead of imposing any arbitrary distribution function, I will evaluate these welfare effects using the empirical distributions of θ and its components.

Nonlinear pricing solutions based on (60) are constructed under the assumption that the monopolist considers only the possibility of allowing for call discounts, instead of price discrimination based on duration of the call, time of the day, distance, or any other criteria.[35] Table 5 presents the results of evaluating these tariffs and their associated welfare effects for the two Kentucky local exchanges where the tariff experiment was conducted. I compute an adaptive Gaussian kernel with optimal bandwidth chosen to minimize the mean integrated square error of the estimation of the distributions of θ, θ_1, and θ_2 (actual or expected calls and estimation bias respectively) corresponding to each local exchange.[36]

Distribution estimates are shown in Figures 3–4 for the two local exchanges of Bowling Green and Kentucky. There are not many differences among the distributions of these exchanges. In both cases, the kernel estimates identify important focal points around 50 and 100 expected calls per week respectively. However, the most important distinction

[35] The available data is not able to identify any effect other than potential volume discounts based on the total number of calls. This is because the database includes expectations for total number of weekly calls during the spring months when the effective marginal tariff is zero, but it does not include anything regarding expected duration of calls or average time/distance profile of these calls.

[36] See Silverman (1986, §5.3). The estimation procedure discretizes the ranges of θ, θ_1, and θ_2 around a 128 point grid to obtain the kernel estimation of each density by means of a fast Fourier transform. Estimation of $f(\cdot)$ and $F(\cdot)$ for intermediate values of θ, θ_1, or θ_2 is obtained by polynomial interpolation (with all 128 point estimates of the kernel) using Neville's algorithm [Press, Flannery, Teulosky, and Vetterling (1986, §3.1)].

between these two exchanges is the positive bias in Louisville. The hazard rate of all distributions can be considered increasing. The first increasing portions of the different hazard rates of Figures 3 and 4 account for most of the corresponding mass of probability. For large θ, θ_1, or θ_2 a small variation of the respective kernel estimate of their probability density functions, $\hat{f}(\cdot)$ will become much larger when divided by their estimated survival functions $1 - \hat{F}(\cdot)$, therefore making the kernel estimate of the hazard rates unreliable for large values of θ, θ_1, or θ_2. As all increasing variances of the kernel estimates of the hazard rates show, most variations in $\hat{r}(\cdot)$ after the initial increasing section are most likely due to purely random effects than to genuine increases or decreases of the hazard rate for particular regions.[37]

Before analyzing the results of the simulations from these kernel estimates of the distributions of types, it is necessary to address two identification issues that will provide with meaningful simulation figures: *a)* What is the actual average number of calls in each local exchange with different tariff regimes when $p \neq 0$? And *b)* What is a reasonable value of the elasticity of demand?

The values of θ and θ_1 are identified as the actual and expected number of calls during the spring months, when consumers faced a zero marginal charge. Thus, the observed and expected consumption are excellent dummies of individual satiation levels that characterize the type of each individual consumer. The existence of a positive charge per call could lead to a selection effect in Louisville where the flat rate was still an option later in the fall, and a suppression effect in Bowling Green (mandatory measured) and Louisville (optional measured). During the fall months in which these tariffs applied, customers in Bowling Green made 134.33 local calls on average every month. This number identifies the number of calls of the *ex–post* tariff in my base case for Bowling Green since it already includes the effect of a positive marginal tariff. In Louisville this number is significantly higher as it averages the number of calls of 30% of the customers on optional measured service, 86.69, and the 189.28 monthly calls of the remaining 70% of customers on optional flat rate service. The value of 179.02 is therefore used in the base case to identify the volume of demand under the *ex–ante* pricing regime in Louisville.

Price elasticity of demand function (60) is given by $\varepsilon = \alpha p$. Therefore, for any price per call, it is always possible to modify the value of α in order to fit a demand function with the desired value of the elasticity. The simulations are run for four values of price elasticity (evaluated at the average sample p) as reported in four independent empirical studies of local telephone demand: -0.1 [Park, Wetzel, and Mitchell (1983)], -0.17 [Kling and Van Der Ploeg (1990)], -0.45 [Train, McFadden, and Ben–Akiva (1987)], and -0.7 [Hobson and Spady (1988)]. Because of the methodology used and the richness of the data available, the estimate of Park, Wetzel, and Mitchell is probably the most accurate. I however decided to choose $\varepsilon = -0.17$ for the base case common to the two cities because this number allows for comparisons with other situations, both with higher and lower elasticities of demands. But

[37] The variance of the hazard rate shown at the bottom of Figures 3–4 has been computed according to equation (6.25) of Silverman (1986, §6.5) for adaptive Gaussian kernels.

also, and more importantly, because Park, Wetzel, and Mitchell (1983, §5) acknowledge that demand elasticities are price dependent and they indicate that elasticities could likely be higher for higher prices (as in my case).

The elasticity is dependent on the price of the calls. Unfortunately the actual tariffs in both Bowling Greene and Louisville were multidimensional, and not only charged for setup but also for duration, and both of them varied depending on the time band, and in some cases the distance of the call within the local exchange. After comparing telephone usage patterns in the two local exchanges, I chose an average cost per call of 7 cents as representative for the base case of the simulations.

Table 5 evaluates each particular nonlinear pricing solution and its associated welfare magnitudes: consumer surplus V, profits (revenues) π, and total welfare W. All average values of simulations in Table 5 are shown in 1986 dollars per month. Reported simulations are the average of 10,000 independent draws from the kernel estimation of the empirical distribution of types.[38] These random draws from the kernel distribution estimates are constrained to have the same first two moments of the true distributions of θ, θ_1, and θ_2 respectively (see Table 1).

To comment results of Table 5, I focus on the case where $\varepsilon = -0.17$. Thus, in Bowling Green, the optimal *ex–post* tariff involves an average marginal rate of \$0.07, and an average monthly fee of \$44.07. Given the empirical distributions of types in that local exchange, consumers enjoy an average expected money surplus of \$11.25, the local monopolist expects to make \$44.92 in profits (revenues) per customer, and total expected welfare amounts to \$56.17 per person.

Average monthly fees are slightly higher under optional pricing than with the standard *ex–post* nonlinear tariffs (SNLT), although almost no distinction is found between optional two–part tariffs (OTPT) and optional nonlinear tariffs (ONLT). Marginal rates are 34% lower with OTPT than with SNLT while under ONLT, they rise 18%. These are however average magnitudes. Thus, the higher consumption under ONLT relative to OTPT could be explained by a likely reduction in marginal tariffs under ONLT relative to OTPT as consumption increases (θ_2) for each chosen tariff (θ_1). This increase in consumption explains the important increase in expected consumer surplus under ONLT. Expected consumer surplus increases by 14% due to the 5% expansion of demand under ONLT relative to SNLT as compared to the 1% expansion with OTPT.

Introduction of OTPT enhance welfare by about 2%, mostly driven by a 4% increase in profits, because consumer surplus is also reduced by 4% (of an initial smaller amount than profits). ONLT reduces welfare by 5%, but the distribution of its components is quite different from the OTPT case. It appears that the effect of the reduction of marginal rates for large consumers under ONLT dominates and thus consumers benefit more from the

[38] The practical advantage of simulation using random draws from kernel estimates of the distributions as compared to simple repeated random draws with replacement from the sample of types is that we can evaluate the pricing functions at any point, even if it has not exactly occurred in the data.

introduction of ONLT than from the introduction of OTPT, although the latter one is the welfare maximizing pricing policy in expectation among the three analyzed here.

Finally, all magnitudes considered (with the exception of consumption) are inversely related to the absolute value of the elasticity of demand. Thus, the more inelastic is the demand, the higher is the average the fixed fee as well as marginal tariffs. But also the expected consumer surplus, profits and total welfare. The analysis of the four different scenarios points out to the importance of an accurate estimate of the elasticity to evaluate welfare effects. For instance, comparing the total expected welfare across different scenarios and tariffs, it increases an average of 70% when $\varepsilon = -0.10$ relative to $\varepsilon = -0.17$. However it falls about 62% when $\varepsilon = -0.45$, and 75% when $\varepsilon = -0.70$. However, the welfare analysis carried out before for the reference scenario when $\varepsilon = -0.17$ is also valid for the others, so that the conclusion of OTPT being the preferred pricing option appears to be robust to different values of the elasticity of demand.

For the case of Louisville, the reference OTPT case is characterized again with an average marginal rate of \$0.07, and the average monthly fee of \$63.59.[39] Individual expected consumer surplus is \$10.03, expected profits per customer are \$65.00, and total expected welfare amounts to \$75.03 per person.

The welfare analysis of the results of Louisville is very similar to that one of Bowling Green. There are two sources of differences between these two exchanges that affect the results of simulations. Consumption pattern may vary due to differences in demographics, socioeconomic variables, and/or the size of the local network. The effect of all these variables have already been captured through the identification of exchange specific levels of telephone usage under different tariff regimes. The other source is the disparate behavior of type shocks in these two cities. Results of Section 5.1 show that θ second order stochastically dominates θ_1 in Bowling Green, while in Louisville, this stochastic dominance is of first order, which means that Louisville customers systematically underestimate their future local telephone usage.

Systematic underestimation of future consumption is the origin of the wider effects of welfare in Louisville relative to Bowling Green when comparing pricing alternatives. Thus, for instance, for the $\varepsilon = -0.17$ scenario, going from SNLT to OTPT reduces the expected consumer rents by 4% and increases expected profits by 4% in Bowling Green, while in Louisville the expected consumer surplus reduction is about 20% and the increase in expected profits reaches 7%. This means a less favorable distribution of welfare for consumers due to the fact that the majority of them would choose particular OTPTs that *ex–post* prove to be designed for very low consumption profiles. However, OTPT is again the welfare maximizing among the three pricing strategies considered here.

[39] Observe that average marginal rates are normalized to \$0.07 both for the OTPT and ONLT cases. Since consumption (independent of ε) is also normalized across scenarios, the average marginal rate is always the same for OTPT and ONLT across scenarios.

Finally, observe that the simulation results regarding differences of expected profits are in accordance to the theoretical results of the previous sections, and of Proposition 6 in particular. Welfare increases in expectation when we implement OTPT instead of SNLT. The SOSD of θ over θ_1 is the dominant factor driving this result. The FOSD of Louisville, with mean increasing effect on the usage level accounts for the stronger effect in the increase of expected profits (7% in Louisville *vs.* 4% in Bowling Green). Finally, the additional 4% increase in profits obtained when ONLT are in use instead of OTPT should be explained by the monopolist being able to discriminate consumers also with respect to θ_2. Expected profits increase as the number of options increases and accounts for *ex–post* differences.

7 Conclusions

This paper has addressed several pricing strategies of a monopolist who wants to screen consumers when their individual demands are stochastic. For tractability considerations, I have restricted my attention to the case where consumer taste parameters remain single–dimensional in order to study the conditions leading to quantity discounts of the optional nonlinear tariffs when consumers buy several units of the same product.

As mentioned in previous sections, optional nonlinear pricing has not attracted much attention among economists until very recently. Traditionally, economists have incorrectly extended the application of results of the standard nonlinear pricing theory to situations where consumption and tariff choice were not simultaneous. The early treatment of Clay, Sibley, and Srinagesh (1992) studied the design of optimal two–part tariffs, but restricting their attention to the discrete type case. They also limited drastically the range of variation of θ_2 to ensure that the same SCP held both *ex–ante* and *ex–post*, so that the ordering of individual consumer preferences remained unaltered after the realization of the shock. Miravete (1996) extended this model to the case of a continuum of two–part tariff options with a continuum of types, independently of whether the ordering of consumer tastes changed or not after the realization of the shock. Miravete (1997) used a particular closed form solution of this model to analyze the estimation bias of not dealing with asymmetric information and self–selection issues in a cross–section framework. Finally, Courty and Li (1998) analyzed a general model of sequential screening with a continuum of types but limiting the analysis to consumers with unit demands and biased type shocks in the sense of FOSD.

Relative to all these works, the present paper contributes by characterizing a fully nonlinear tariff when consumers buy more than one unit, and by making explicit the role of the statistical assumptions on the existence of quantity discounts (IHR of the distribution of type components), and welfare effects (FOSD and SOSD of θ over θ_1). This paper also compares different optimal nonlinear tariffs depending on whether they are designed *ex–ante* or *ex–post*, through the preservation of the IHR property of the distribution of type components through convolution. Finally, the paper also contributes to this literature by providing very strong evidence in favor of the suggested type–varying model based on direct observation of consumer types. Furthermore, using simulations from the kernel distributions of these types, the paper reports results that favor optional two–part tariffs as the welfare maximizing strategy in two local exchanges of Kentucky.

The results of this paper help explaining the widespread use of tariff discounts embodied into tariff options in several monopolistic and competitive markets. But the suggested solution also opens the possibility of theoretical extensions to other agency problems where individual stochastic components of moral hazard or adverse selection parameters could also be present. Restricting our attention to common nonlinear pricing issues, there are no significant difficulties (besides burdensome notation) in extending the present model to address for example the case of firms that compete through the design of optional nonlinear tariffs. This could be done along the line of the papers by Armstrong and Vickers (1998) or Stole (1995) for the case of differentiated products, or Rochet and Stole (1998) when the competing firms sell an homogeneous good. Extensions to Principal–Agent problems based on Laffont and Tirole's (1993) model could also open the possibility of analyzing additional features associated to the existence of uncertainty in several issues with direct policy applications, both from the perspective of adverse selection and moral hazard.

References

Anderson, G. (1996): "Tests of Stochastic Dominance in Income Distributions." *Econometrica*, 64, 1183–1194.

Armstrong, M. (1996): "Multiproduct Nonlinear Pricing." *Econometrica*, 64, 51–75.

Armstrong, M. and J. Vickers (1998): "Competitive Price Discrimination." Mimeo, Oxford University.

Bagnoli, M. and T. Bergstrom (1989): "Log–Concave Probability and Its Applications." University of Michigan Center for Research on Economic and Social Theory Working Paper #89–23.

Barlow, R.E., A.W. Marshall, and F. Proschan (1963): "Properties of Probability Distributions with Monotone Hazard Rate." *Annals of Mathematical Statistics*, 34, 375–389.

Barlow, R.E. and F. Proschan (1975): *Statistical Theory of Reliability and Life Testing*. Holt, Rinehart, and Winston, Inc.

Baron, D.P. and D. Besanko (1984): "Regulation and Information in a Continuing Relationship." *Information Economics and Policy*, 1, 267–302.

Brown, S.J. and D.S. Sibley (1986): *The Theory of Public Utility Pricing*. Cambridge University Press.

Caillaud, B., R. Guesnerie, and P. Rey (1992): "Noisy Observation in Adverse Selection Models." *Review of Economic Studies*, 59, 595–615.

Clay, K., D.S. Sibley, and P. Srinagesh (1992): "Ex Post *vs.* Ex Ante Pricing: Optional Calling Plans and Tapered Tariff." *Journal of Regulatory Economics*, 4, 115–138.

Courty, P. and H. Li (1998): "Sequential Screening." Mimeo, Universitat Pompeu Fabra.

Eaton, M.L. and M.D. Perlman (1991): "Multivariate Probability Inequalities: Convolution Theorems, Composition Theorems, and Concentration Inequalities," in K. Mosler and M. Scarsini (eds.): *Stochastic Orders and Decision Under Risk*. Institute of Mathematical Statistics.

Faulhaber, G.R. and J.C. Panzar (1977): "Optimal Two–Part Tariffs with Self–Selection." Bell Laboratories Economic Discussion Paper No. 74.

Gantmacher, F.R. (1959): *Matrix Theory*. Chelsea Publishing Company.

Goldman, M.B., H.E. Leland, and D.S. Sibley (1984): "Optimal Nonuniform Prices." *Review of Economic Studies*, 51, 305–319.

Hadar, J., and W.R. Russell (1969): "Rules for Ordering Uncertain Prospects." *American Economic Review*, 59 25–34.

Hogg, R.V. and A.T. Craig (1995): *Introduction to Mathematical Statistics*, 5th edition. Macmillan Publishing Co., Inc.

Hobson, M. and R.H. Spady (1988): "The Demand for Local Telephone Service Under Optional Local Measured Service." Bellcore Economics Discussion Paper No. 50.

Ivaldi, M. and D. Martimort (1994): "Competition under Nonlinear Pricing." *Annales d'Economie et de Statistique*, 34, 71–114.

Johnson, N.L., S. Kotz, and N. Balakrishnan (1995): *Continuous Univariate Distributions*, 2nd edition. John Wiley & Sons.

Kamien, M.I., and N.L. Schwartz (1991): *Dynamic Optimization*, 2nd edition. North–Holland.

Karlin, S. (1968): *Total Positivity*, Vol. I. Stanford University Press.

Kling, J.P. and S.S. van der Ploeg (1990): "Estimating Local Call Elasticities with a Model of Stochastic Class of Service and Usage Choice," in A. de Fontenay, M.H. Shugard, and D.S. Sibley (eds.): *Telecommunications Demand Modelling*. North–Holland.

Kotlarski, I. (1962): "On Groups of n Independent Random Variables Whose Product Follows the Beta Distribution." *Colloquium Mathematicum*, 9, 325–332.

Kridel, D., D. Lehman, and D. Weisman (1993): "Option Value, Telecommunications Demand, and Policy." *Information Economics and Policy*, 17, 69–75.

Laffont, J.J. (1993): *The Economics of Uncertainty and Information*. MIT Press.

Laffont, J.J., H. Ossard, and Q. Vuong (1995): "Econometrics of First Price Auctions," *Econometrica*, 63, 953–980.

Laffont, J.J. and J. Tirole (1986): "Using Cost Observations to Regulate Firms." *Journal of Political Economy*, 94, 614–641.

Laffont, J.J. and J. Tirole (1993): *A Theory of Incentives in Procurement and Regulation*. MIT Press.

MacKie–Mason, J.K. and D. Lawson (1993): "Local Telephone Calling Demand when Customers Face Optimal and Nonlinear Price Schedules." Working Paper. Department of Economics. University of Michigan.

Marshall, A.W. and I. Olkin (1979): *Inequalities: Theory of Majorization and Its Applications*. Academic Press.

Mas–Colell, A., M.D. Whinston, and J.R. Green (1995): *Microeconomic Theory*. Oxford University Press.

Maskin, E. and J. Riley (1984): "Monopoly with Incomplete Information." *Rand Journal of Economics*, 15, 171–196.

McDonald, J.B. and Y.J. Xu (1995) "A Generalization of the Beta Distribution with Applications." *Journal of Econometrics*, 66, 133–152; plus "Errata," 69, 427–428.

Miravete, E.J. (1996): "Screening Consumers Through Alternative Pricing Mechanisms." *Journal of Regulatory Economics*, 9, 111–132.

Miravete, E.J. (1997): "Estimating Demand for Local Telephone Service with Asymmertric Information and Optional Calling Plans." INSEAD Working Paper 97/43/EPS.

Mitchell, B.M. and I. Vogelsang (1991): *Telecommunications Pricing. Theory and Practice*. Cambridge University Press.

Ng, Y.K. and M. Weisser (1974): "Optimal Pricing with a Budget Constraint: The Case of the Two–Part Tariff." *Review of Economic Studies*, 41, 337–345.

Oren, S.S., S.A. Smith, and R.B. Wilson (1985): "Capacity Pricing." *Econometrica*, 53, 545–566.

Panzar, J.C. and D.S. Sibley (1978): "Public Utility Pricing under Risk: The Case of Self–Rationing." *American Economic Review*, 68, 887–895.

Park, R.E., B.M. Wetzel, and B.M. Mitchell (1983): "Price Elasticities for Local Telephone Calls." *Econometrica*, 51, 1699–1730.

Pečarič, J.E., F. Proschan, and Y.L. Tong (1992): *Convex Functions, Partial Orderings, and Statistical Applications*. Academic Press.

Polyanin, A.D. and A.V. Manzhirov (1998): *Handbook of Integral Equations*. CRC Press.

Prékopa, A. (1971): "Logarithmic Concave Measures with Application to Stochastic Programming." *Acta Scientiarum Mathematicarum*, 32, 301–316.

Press, W., B. Flannery, S. Teulosky, and W. Vetterling (1986): *Numerical Recipes in C: The Art of Scientific Computing*, 2nd edition. Cambridge University Press.

Rochet, J.C. and P. Choné (1998): "Ironing, Sweeping and Multidimensional Screening." *Econometrica*, 66, 783–826.

Rochet, J.C. and L.A. Stole (1998): "Selection Contracts with Uncertain Participation: Re–examining Monopoly and Competitive Nonlinear Pricing." Mimeo, IDEI, Université des Sciences Sociales, Toulouse.

Rogerson, W.P. (1988): "On the Optimality of Menus of Linear Contracts." The Center for Mathematical Studies in Economics and Management Science, Discussion Paper # 714R. Northwestern University.

Silverman, B.W. (1986): *Density Estimation for Statistics and Data Analysis.* Chapman & Hall/CRC.

Springer, M.D. (1979): *The Algebra of Random Variables.* John Wiley & Sons.

Spulber, D.F. (1992): "Optimal Nonlinear Pricing and Contingent Contracts." *International Economic Review*, 33, 747–772.

Srinagesh, P. (1992): "A Dynamic Stochastic Model of Choice." Bellcore Economics Discussion Paper No. 78.

Stole, L.A. (1995): "Nonlinear Pricing and Oligopoly." *Journal of Economic Management and Strategy*, 4, 529–562.

Stoline, M.R. and H.K. Ury (1979): "Tables of the Studentized Maximum Modulus Distribution and an Application to Multiple Comparisons Among Means." *Technometrics*, 21, 87–93.

Taylor, L.D. (1994): *Telecommunications Demand in Theory and Practice*, 2nd edition. Kluwer Academic Publishers.

Tirole, J. (1989): *The Theory of Industrial Organization.* MIT Press.

Train, K.E., D.L. McFadden, and M. Ben–Akiva (1987): "The Demand for Local Telephone Service: A Fully Discrete Model of Residential Calling Patterns and Service Choices." *Rand Journal of Economics*, 18, 109–123.

Train, K.E., M. Ben–Akiva, and T. Atherton (1989): "Consumption Patterns and Self–Selecting Tariffs." *The Review of Economics and Statistics*, 50, 62–73.

Wilson, R.B. (1993): *Nonlinear Pricing.* Oxford University Press.

Wilson, R.B. (1995): "Nonlinear Pricing and Mechanism Design," in H. Amman, D. Kendrick, and J. Rust (eds.): *Handbook of Computational Economics*, Vol. I. North–Holland.

Wolak, F. (1996): "Estimating Regulated Firm Production Functions with Private Information: An Application to California Water Utilities," *Annales d'Economie et de Statistique*, 34, 13–69.

Appendix 1

• Derivation of the Ex–Post Tariff

The corresponding Hamiltonian for the monopolist's ex–post problem is:

$$H[V,p,\theta] = [v(p(\theta),\theta) - V(\theta) - (p(\theta)-c)v_p(p,\theta) - K]\,f(\theta) + \lambda(\theta)v_\theta(p(\theta),\theta) \qquad (A.1)$$

Using equation (2), the first order necessary conditions are:

$$H_p : -(p(\theta)-c)v_{pp}(p,\theta)f(\theta) + \lambda(\theta)v_{p\theta}(p(\theta),\theta) = 0 \qquad (A.2)$$

$$H_V : f(\theta) = \lambda'(\theta) \qquad ; \qquad \lambda(\overline{\theta}) = 0 \qquad (A.3)$$

There is not transversality condition at $\overline{\theta}$ since $V'(\overline{\theta}) > 0$ because the participation constraint is only binding at $\underline{\theta}$. Then [Kamien and Schwartz (1991, §II.7)]:

$$\lambda(\theta) = \int_{\overline{\theta}}^{\theta} f(z)dz = F(\theta) - 1 \qquad (A.4)$$

Equations (9) – (10) follow from substituting this expression and the SCP into the first order necessary conditions H_p and H_V.

• Proof of Lemma 1

Assumption 3 restricts the distribution functions considered here to be continuously differentiable and positive. From Definiton 4 it follows that a twice continuously differentiable function $g(z)$ is log–concave if and only if:

$$\frac{\partial^2[\log g(z)]}{\partial z^2} = \frac{\partial}{\partial z}\left[\frac{g'(z)}{g(z)}\right] < 0, \qquad (A.5)$$

i.e., since $g(z) > 0\ \forall z \in \Re$, $g'(z)/g(z)$ has to be monotone decreasing on R. Next, without loss of generality, assume $x_1 < x_2$ and $0 = y_1 < y_2 = \Delta$. Then, from the definition of PF_2 the following inequality should hold:

$$g(x_1)g(x_2-\Delta) - g(x_2)g(x_1-\Delta) \geq 0. \qquad (A.6)$$

Based on this inequality, all the following inequalities are equivalent:

$$\frac{g(x_2-\Delta)}{g(x_2)} \geq \frac{g(x_1-\Delta)}{g(x_1)}, \qquad (A.7a)$$

$$\lim_{\Delta\to 0}\frac{g(x_2-\Delta)}{g(x_2)} \geq \lim_{\Delta\to 0}\frac{g(x_1-\Delta)}{g(x_1)}, \qquad (A.7b)$$

$$\frac{-g'(x_2)}{g(x_2)} \geq \frac{-g'(x_1)}{g(x_1)}, \qquad (A.7c)$$

which, given $g(z) > 0$, proves that $\forall z \in \Re$, $g'(z)/g(z)$ is monotone decreasing on $\Re$. ■

• **Proof of Lemma 2**

By definition of PF_2, the convolution distribution $F(\theta)$ defined in (46) has to be such that $\forall x_1, x_2 \in X \subseteq \Re$ and $\forall y_1, y_2 \in Y \subseteq \Re$, such that $x_1 < x_2$ and $y_1 < y_2$, the following condition holds:

$$\begin{vmatrix} F(x_1 - y_1) & F(x_1 - y_2) \\ F(x_2 - y_1) & F(x_2 - y_2) \end{vmatrix} = \begin{vmatrix} \int\limits_R F_1(x_1 - z) f_2(z - y_1) dz & \int\limits_R F_1(x_1 - z) f_2(z - y_2) dz \\ \int\limits_R F_1(x_2 - z) f_2(z - y_1) dz & \int\limits_R F_1(x_2 - z) f_2(z - y_2) dz \end{vmatrix}$$

$$= \int\limits_{z_1 <} \int\limits_{z_2} \begin{vmatrix} F_1(x_1 - z_1) & F_1(x_1 - z_2) \\ F_1(x_2 - z_1) & F_1(x_2 - z_2) \end{vmatrix} \cdot \begin{vmatrix} f_2(z_1 - y_1) & f_2(z_1 - y_2) \\ f_2(z_2 - y_1) & f_2(z_2 - y_2) \end{vmatrix} dz_1 dz_2 \geq 0, \qquad (A.8)$$

where the last inequality is the *Basic Composition Formula* that relates convolutions of totally positive functions. From here the proof is immediate since the first determinant in the double integral is positive as $F_1(\theta_1)$ is PF_2 and the second determinant is also positive as $f_2(\theta_2)$ is PF_2. Indices can be changed arbitrarily and the convolution be referred to other functions (density, survival) as long as they fulfill the assumptions stated in this lemma. ■

• **Proof of Proposition 3**

In order to prove this proposition let first study the total positivity properties of the function $\delta : \Re \rightarrow \{0, 1\}$ defined as follows:

$$\delta(x - y) = \begin{cases} 0 & \text{if} \quad x < y \\ 1 & \text{otherwise} \end{cases} \qquad (A.9)$$

From Definition 5, $\delta(x - y)$ is PF_2 if $\forall x_1, x_2 \in X \subseteq R$ and $\forall y_1, y_2 \in Y \subseteq \Re$, such that $x_1 < x_2$ and $y_1 < y_2$, the following condition holds:

$$\begin{vmatrix} \delta(x_1 - y_1) & \delta(x_1 - y_2) \\ \delta(x_2 - y_1) & \delta(x_2 - y_2) \end{vmatrix} \geq 0. \qquad (A.10)$$

There are six possible cases:

1. If $x_1 < x_2 < y_1 < y_2$, then $0 \cdot 0 - 0 \cdot 0 = 0$,
2. If $x_1 < y_1 \leq x_2 < y_2$, then $0 \cdot 0 - 1 \cdot 0 = 0$,
3. If $x_1 < y_1 < y_2 \leq x_2$, then $0 \cdot 1 - 1 \cdot 0 = 0$,
4. If $y_1 \leq x_1 < y_2 \leq x_2$, then $1 \cdot 1 - 1 \cdot 0 = 1$,
5. If $y_1 \leq x_1 < x_2 < y_2$, then $1 \cdot 0 - 1 \cdot 0 = 0$,
6. If $y_1 < y_2 \leq x_1 < x_2$, then $1 \cdot 1 - 1 \cdot 1 = 0$.

Thus $\delta(x-y)$ is PF_2. It is then straightforward to show that $\hat{\delta}(x-y) = 1 - \delta(x-y)$ is also PF_2. By Lemma 2, $\hat{\gamma}(\theta_i)$, the convolution of $\hat{\delta}(x-\theta_i)$ and $f_i(\theta_i)$ is PF_2. Therefore:

$$\hat{\gamma}(\theta_i) = \int\limits_R \hat{\delta}(x-\theta_i) f_i(\theta_i) d\theta_i = \int\limits_{-\infty}^{x} f_i(\theta_i) d\theta_i, \tag{A.11}$$

because $\hat{\delta}(x-\theta_i) = 1$ only if $x < \theta_i$. Hence, in this case, $\hat{\gamma}(\theta_i) = F_i(\theta_i = x)$, and therefore the cumulative distribution function $F_i(\theta_i)$ is PF_2. Similarly, $\gamma(\theta_i)$ the convolution of $\delta(x-\theta_i)$ and $f_i(\theta_i)$ is also PF_2, which in this case implies that:

$$\gamma(\theta_i) = \int\limits_R \delta(x-\theta_i) f_i(\theta_i) d\theta_i = \int\limits_{x}^{\infty} f_i(\theta_i) d\theta_i, \tag{A.12}$$

because $\delta(x-\theta_i) = 1$ only if $x \geq \theta_i$. Thus, in this second case, $\gamma(\theta_i) = \overline{F}_i(\theta_i = x)$, and therefore the survival function $1 - F_i(\theta_i)$ is also PF_2. ∎

• Proof of Proposition 4

Suppose not, *i.e.*, for instance assume that $r_1(\theta) < r(\theta)$:

$$\frac{f_1(\theta)}{\overline{F}_1(\theta)} < \frac{f(\theta)}{\overline{F}(\theta)}. \tag{A.13}$$

Using the definition of Fourier convolution, this inequality is equivalent to the following three inequalities:

$$f_1(\theta)\overline{F}(\theta) - f(\theta)\overline{F}_1(\theta) < 0, \tag{A.14}$$

$$f_1(\theta) \int\limits_0^{\infty} \overline{F}_1(\theta - z) f_2(z) dz - \overline{F}_1(\theta) \int\limits_0^{\infty} f_1(\theta - z) f_2(z) dz < 0, \tag{A.15}$$

$$\int\limits_0^{\infty} \left[f_1(\theta)\overline{F}_1(\theta - z) - \overline{F}_1(\theta) f_1(\theta - z) \right] f_2(z) dz < 0. \tag{A.16}$$

Since by Assumption 2, $f_2(\theta) \geq 0$ on $0 \leq \theta < \infty$, it must be the case that the term between brackets is negative $\forall \theta \geq 0$. But observe that this condition then requires

$$\frac{f_1(\theta)}{\overline{F}_1(\theta)} \leq \frac{f_1(\theta - z)}{\overline{F}_1(\theta - z)} \qquad \forall z \geq 0 \tag{A.17}$$

so that $F_1(\theta_1)$ should be decreasing hazard rate. Similarly, $r_2(\theta) < r(\theta)$ violates $F_2(\theta_2)$ being IHR. Contradiction. ∎

• **Tariff Solutions for Exponential Demand**

I present here the solutions of the optimal *ex–post* tariff, the *ex–ante* optional two–part tariffs, and the *ex–ante* nonlinear options. These solutions are constructed for demand specification (60). These expressions are then used to evaluate the corresponding welfare magnitudes. For simplicity I have assumed that $c = 0$ and $K = 0$. These parameters only change the scale of the fixed fee and marginal charge, but the comparisons of expected welfare under pricing regimes remain unaffected. Thus, the optimal *ex–post* tariff is characterized by:

$$\hat{p}(\theta) = \frac{1}{\alpha\theta r(\theta)}, \tag{18a}$$

$$\hat{A}(\theta) = \frac{1}{\alpha}\left\{\theta \exp[-\{\theta r(\theta)\}^{-1}] - \int_{\underline{\theta}}^{\theta} \exp[-\{zr(z)\}^{-1}]dz\right\}. \tag{18b}$$

The expressions for the optional two–part tariffs are more cumbersome. However, for the present application the sample only includes active consumers. Thus, the *ex–post* participation constraints are not binding for any *ex–ante* type. This fact simplifies optimal pricing expressions and computations because $F_2[\underline{\theta}_2(\theta_1)] = 0$ for all possible *ex–ante* types θ_1, and because $E_2[\theta_2 \mid \theta_2 \geq \underline{\theta}_2(\theta_1)] = \mu_2$, which is straightforward to compute from the available data. Taking advantage of these simplifications, the menu of optional two–part tariffs is given by:

$$\tilde{p}(\theta_1) = \frac{1}{\alpha(\theta_1 + \mu_2)r_1(\theta_1)}, \tag{19a}$$

$$\tilde{A}(\theta_1) = \frac{\exp[-\{(\theta_1 + \mu_2)r_1(\theta_1)\}^{-1}]}{\alpha}\left\{(\theta_1 + \mu_2) - \int_{\underline{\theta}_1}^{\theta_1} \exp[-\{(z + \mu_2)r_1(z)\}^{-1}]dz\right\}. \tag{19b}$$

There are however some difficulties because since the shock is defined as the difference between the actual and expected telephone usage, its empirical support is dependent on the magnitude of the expectation. Since only active customers are included in the sample, a consumer that expects to make 15 weekly calls cannot suffer a negative demand shock higher than 15 because otherwise her actual consumption would be negative, which makes no sense. Nevertheless, simulations can handle all these issues by imposing that the random draws fulfill this non–negativity condition, which appears to bind only for a small percentage of the sample.

The expressions for the optional nonlinear options are further complicated by the existence of $\theta_2^*(\theta_1)$ in the solutions that characterize the optimal nonlinear tariff options. Fortunately, it is straightforward to show that for the utility function (59), condition (31) only requires that $\theta_2^*(\theta_1) = \mu_2$ for all possible *ex–ante* types θ_1. The optimal departure from the "boundary two–part tariff" is:

$$\Delta\tilde{\tilde{p}}(\theta) = \frac{F_2(\mu_2) - F_2(\theta_2)}{\alpha(\theta_1+\theta_2)f_2(\theta_2)}, \qquad (20a)$$

$$\Delta\tilde{\tilde{A}}(\theta) = \frac{\exp\left[-\{(\theta_1+\mu_2)r_1(\theta_1)\}^{-1}\right]}{\alpha}\Bigg\{(\theta_1+\theta_2)\left(\exp\left[\frac{F_2(\mu_2)-F_2(\theta_2)}{(\theta_1+\theta_2)f_2(\theta_2)}\right]-1\right)$$

$$-\int\limits_{\mu_2}^{\theta_2}\left(\exp\left[\frac{F_2(\mu_2)-F_2(z)}{(\theta_1+z)f_2(z)}\right]-1\right)dz\Bigg\}, \qquad (20b)$$

which together with $(19a)$ – $(19b)$ defines the menu of nonlinear options $\{\tilde{\tilde{A}}(\theta), \tilde{\tilde{p}}(\theta)\}$.

Appendix 2

• Description of Variables

The data set includes the following variables. Most of them are dummies that take value equal to 1 for the indicated case:

CALLS	Monthly average of weekly number of actual calls.
EXPCALLS	Expected number of weekly calls.
BIAS	*Calls – Expcalls.*
AGE1	The head of the household is between 15 and 34 years old.
AGE2	The head of the household is between 35 and 54 years old.
AGE3	The head of the household is at least 54 years old.
BENEFITS	The household receives some benefits such as Food Stamps, Social Security, Federal Rent Assistance, Aid to Families with Dependent Children,...
BLACK	The head of the household belongs to the black ethnic group.
CHURCH	Some member of the household uses the telephone for charity or church work.
COLLEGE	The head of the household is at least a college graduate.
HHSIZE	Number of people who regularly live in the household.
INCOME	Estimated total monthly income of the household.
DINCOME	The household did not report its annual income. Recoded at $19,851.
MARRIED	The head of the household is married.
MEASURED	The household is on local measured service in one particular month.
MOVED	The household moved at least once in the last five years.
ONLYMALE	The head of the household is single and male.
RETIRED	The head of the household is retired.
TEENS	Number of teenagers (between 13 and 19 years old) living in the household.
MARCH	Observation for the month of March 1986.
APRIL	Observation for the month of April 1986.
MAY	Observation for the month of May 1986.
OCTOBER	Observation for the month of October 1986.
NOVEMBER	Observation for the month of November 1986.
DECEMBER	Observation for the month of December 1986.

Appendix 3

• Test of Stochastic Dominance

In this Appendix I will follow the notation of Anderson (1996) to identify the elements of the stochastic dominance test referred to the model of this paper. The partitions of the range of phone are somewhat arbitrary and have been chosen to define fractiles that are more or less homogeneous in size. The definition of k fractiles is complicated by the existence of "focal points" in the distribution of expected phone calls, which is evident from Figures 1.1 and 1.2. These category thresholds define the interval length d_j, and each distribution determines the number of cases $x(\theta)$, $x(\theta_1)$ within each category. Given that the number of observations (5,241 in Bowling Green and 4,349 in Louisville) is the same for the actual and expected number of calls ($n(\theta) = n(\theta_1) = n$), it is straightforward to obtain the probability of any number of calls falling in a given category under each distribution, $\phi(\theta) = x(\theta)/n(\theta)$, $\phi(\theta_1) = x(\theta_1)/n(\theta_1)$. The test makes use of $p^\star = (x(\theta) + x(\theta_1))/(n(\theta) + n(\theta_1))$, the probability of the joint sample of actual and expected calls, to build the covariance matrix and test the significance of the differences referred to (52) and (53). It is useful to define two matrices to state the FOSD and SOSD hypotheses:

$$I_f = \begin{pmatrix} 1 & 0 & 0 & \cdot & \cdot & 0 \\ 1 & 1 & 0 & \cdot & \cdot & 0 \\ 1 & 1 & 1 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ 1 & 1 & 1 & \cdot & \cdot & 1 \end{pmatrix} \;;\; I_F = \frac{1}{2} \begin{pmatrix} d_1 & 0 & 0 & \cdot & \cdot & 0 \\ d_1 + d_2 & d_2 & 0 & \cdot & \cdot & 0 \\ d_1 + d_2 & d_2 + d_3 & d_3 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ d_1 + d_2 & d_2 + d_3 & d_3 + d_4 & \cdot & \cdot & d_k \end{pmatrix}. \quad (A.21)$$

Thus, the FOSD hypotheses can be written as:

$$H_0 : I_f(\phi(\theta) - \phi(\theta_1)) = 0 \qquad vs. \qquad H_1 : I_f(\phi(\theta) - \phi(\theta_1)) \leq 0, \quad (A.22)$$

and the SOSD hypotheses as:

$$H_0 : I_F I_f(\phi(\theta) - \phi(\theta_1)) = 0 \qquad vs. \qquad H_1 : I_F I_f(\phi(\theta) - \phi(\theta_1)) \leq 0. \quad (A.23)$$

Given the covariance matrix:

$$\frac{\Omega}{n} = \begin{pmatrix} \phi_1^\star(1 - \phi_1^\star) & -\phi_1^\star\phi_2^\star & \cdots & -\phi_1^\star\phi_k^\star \\ -\phi_2^\star\phi_1^\star & \phi_2^\star(1 - \phi_2^\star) & \cdots & -\phi_2^\star\phi_k^\star \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ -\phi_k^\star\phi_1^\star & -\phi_k^\star\phi_2^\star & \cdots & \phi_k^\star(1 - \phi_k^\star) \end{pmatrix}, \quad (A.24)$$

and $m = n^{-1}(n(\theta) + n(\theta_1))/n(\theta)n(\theta_1)$, the vector of frequency differences, $\phi(\theta) - \phi(\theta_1)$ is normally distributed as $N(0, m\Omega)$. Therefore, for the FOSD hypothesis $I_f(\phi(\theta) - \phi(\theta_1))$ is normally distributed as $N(0, mI_f\Omega I_f')$; and for the SOSD case, $I_F I_f(\phi(\theta) - \phi(\theta_1))$ is normally distributed as $N(0, mI_F I_f \Omega I_f' I_F')$. Tables 3 and 4 present the division of each element, $I_f(\phi(\theta) - \phi(\theta_1))$ and $I_F I_f(\phi(\theta) - \phi(\theta_1))$ by its appropriate standard deviation. Since we need to compare multiple differences simultaneously, the critical values for these ratios are given by the studentized maximum modulus distribution [Stoline and Ury (1979)]. If any of these ratios is significantly positive for any given fractile, the corresponding hypothesis of stochastic dominance should be rejected.

Appendix 4

Table 1. Descriptive Statistics

	Bowling Green	Louisville	TEST
CALLS, θ	32.0489	36.6112	-6.63
	(26.902)	(38.197)	
EXPCALLS, θ_1	31.4137	25.9329	8.02
	(36.123)	(30.827)	
BIAS, θ_2	0.6352	10.6783	-12.64
	(37.179)	(39.966)	
log(INCOME)	7.3097	7.0847	13.55
	(0.798)	(0.819)	
HHSIZE	2.7960	2.5381	9.02
	(1.266)	(1.493)	
TEENS	0.3711	0.2309	10.31
	(0.713)	(0.619)	
DINCOME	0.1328	0.1603	-3.78
	(0.339)	(0.370)	
AGE1	0.0614	0.0625	-0.22
	(0.240)	(0.242)	
AGE2	0.2524	0.2644	-1.34
	(0.434)	(0.441)	
AGE3	0.6861	0.6730	1.37
	(0.464)	(0.469)	
COLLEGE	0.2803	0.2244	6.31
	(0.449)	(0.417)	
MARRIED	0.6926	0.5059	18.85
	(0.462)	(0.500)	
RETIRED	0.1525	0.2550	-12.40
	(0.360)	(0.436)	
BLACK	0.0622	0.1168	-9.25
	(0.242)	(0.321)	
CHURCH	0.2082	0.1692	4.88
	(0.406)	(0.375)	
BENEFITS	0.2063	0.3152	-12.11
	(0.405)	(0.465)	
MOVED	0.4820	0.4074	7.34
	(0.500)	(0.491)	
ONLYMALE	0.0452	0.1053	-10.99
	(0.208)	(0.307)	
MARCH	0.3288	0.3325	-0.38
	(0.470)	(0.471)	
APRIL	0.3318	0.3318	0.00
	(0.471)	(0.471)	
MAY	0.3394	0.3357	0.38
	(0.474)	(0.472)	
Observations	5241	4349	

Mean and standard deviations (between parentheses) of demographics for the spring sample. The "TEST" column shows the test of differences of means for each variable in these two cities.

Table 2. Consumption Expectation Bias

Bowling Green					Louisville			
Obs.	PAT	Avg.Bias	Std.Dev.	Strata	Avg.Bias	Std.Dev.	PAT	Obs.
5241	2652.59	0.6352	(37.179)	ALL	10.6783	(39.966)	2353.89	4249
1723	879.39	0.9765	(37.076)	MARCH	11.6001	(43.581)	758.78	1446
1739	903.94	0.6571	(37.014)	APRIL	10.5580	(39.119)	791.41	1443
1779	879.94	0.2834	(37.457)	MAY	9.8842	(36.946)	819.24	1460
1967	1029.82	2.9062	(39.662)	LOW INCOME	15.9668	(50.592)	917.78	1645
3274	1662.00	-0.7291	(35.541)	HIGH INCOME	7.4610	(31.388)	1484.04	2704
714	293.15	0.0920	(18.198)	HHSIZE=1	6.2131	(34.470)	597.57	1095
1774	1016.19	-1.1249	(30.470)	HHSIZE=2	6.4538	(27.637)	874.67	1502
1290	704.12	2.9518	(33.353)	HHSIZE=3	13.8281	(38.995)	426.18	776
980	562.48	-0.0021	(47.312)	HHSIZE=4	14.3265	(43.909)	336.77	582
483	281.00	3.0087	(59.734)	HHSIZE $\geq$ 5	27.6001	(71.748)	277.91	394
3798	1941.58	-0.3655	(29.838)	TEENS=0	7.5578	(35.786)	2060.40	3653
1029	611.62	0.9405	(54.873)	TEENS=1	23.4185	(47.131)	252.33	460
414	225.09	9.0571	(42.156)	TEENS $\geq$ 2	34.1479	(65.503)	164.79	236
322	217.03	-4.7589	(26.910)	AGE1=1	8.4026	(32.578)	205.51	272
1323	869.76	-2.7377	(42.171)	AGE2=1	9.0469	(38.949)	723.88	1150
3596	1677.65	2.3592	(35.866)	AGE3=1	11.5307	(40.955)	1514.95	2927
1469	828.09	-3.4543	(37.277)	COLLEGE=1	4.6580	(28.899)	524.11	976
3772	1878.68	2.2279	(37.024)	COLLEGE=0	12.4203	(42.480)	1908.92	3373
3630	1851.96	0.5463	(36.427)	MARRIED=1	10.6344	(32.603)	1243.15	2200
1611	835.40	0.8355	(38.830)	MARRIED=0	10.7232	(46.315)	1166.71	2149
799	338.42	1.3146	(28.672)	RETIRED=1	9.6512	(35.496)	561.92	1109
4442	2361.63	0.5130	(38.512)	RETIRED=0	11.0299	(41.384)	1844.82	3240
326	237.93	11.6811	(71.411)	BLACK=1	29.3614	(66.110)	454.15	508
4915	2488.20	-0.0974	(33.587)	BLACK=0	8.2073	(34.340)	1957.76	3841
1091	600.92	-1.8867	(45.088)	CHURCH=1	7.8696	(52.922)	329.06	736
4150	2107.23	1.2982	(34.779)	CHURCH=0	11.2505	(36.754)	2056.26	3613
1081	493.97	2.2926	(35.188)	BENEFITS=1	13.8292	(42.011)	726.25	1371
4160	2201.68	0.2046	(37.671)	BENEFITS=0	9.2277	(38.910)	1661.81	2978
2526	1334.84	0.0820	(40.646)	MOVED=1	10.7220	(39.305)	1100.09	1772
2715	1381.03	1.1500	(33.634)	MOVED=0	10.6482	(40.422)	1303.97	2577
237	145.27	-3.5797	(23.912)	ONLYMALE=1	4.6319	(27.237)	265.54	458
5004	2541.78	0.8349	(37.682)	ONLYMALE=0	11.3900	(41.151)	2127.43	3891

"PAT" column reports Pearson analog goodness of fit test for equality of the distribution "of the expected and actual number of calls. This test is distributed as a $\chi^2(19)$, with 0.05 and 0.01" critical values at 30.14 and 36.19 respectively. All statistics have p–values lower than 0.01.

Table 3a. Tests of First Order Stochastic Dominance. Bowling Green

Call Range:	0–3	4–5	6–7	8–9	10	11–12	13–14	15	16–19	20	21–22	23–25	26–29	30–31	32–35	35–44	45–50	51–60	61–90	>90
ALL	-6.27	-5.18	-2.32	2.01	-8.47	-6.27	-1.09	-9.95	0.64	-10.95	-7.68	-10.64	-3.82	-10.71	-8.26	-2.46	-9.23	-3.90	2.72	0.00
MARCH	-3.02	-2.85	-1.57	0.74	-5.35	-3.69	-1.17	-6.12	0.21	-6.54	-4.64	-6.58	-2.50	-6.28	-4.52	-1.71	-6.00	-2.95	1.03	0.00
APRIL	-4.56	-3.00	-1.04	1.41	-4.81	-3.93	-0.31	-5.46	0.38	-6.47	-4.41	-5.74	-2.14	-6.33	-5.26	-1.71	-5.33	-2.05	1.52	0.00
MAY	-3.36	-3.11	-1.41	1.32	-4.53	-3.25	-0.42	-5.65	0.52	-5.96	-4.26	-6.11	-1.98	-5.95	-4.52	-0.85	-4.67	-1.75	2.16	0.00
LOW INCOME	-5.00	-5.21	-4.51	-1.55	-7.11	-6.42	-3.16	-8.64	-1.85	-8.46	-6.73	-8.31	-3.86	-8.24	-7.00	-3.58	-8.19	-5.16	-0.08	0.00
HIGH INCOME	-3.94	-2.46	0.73	3.91	-5.16	-2.89	1.14	-5.85	2.28	-7.29	-4.51	-7.03	-1.86	-7.20	-5.06	-0.41	-5.46	-1.06	3.43	0.00
HHSIZE=1	-0.72	-0.58	1.09	2.64	-2.21	-2.03	0.32	-3.41	0.11	-2.07	-0.40	-2.11	0.33	-0.50	-0.63	0.09	-1.76	-1.28	-0.28	0.00
HHSIZE=2	-6.28	-3.39	-2.11	1.68	-4.90	-3.09	0.64	-4.58	2.89	-5.21	-3.33	-4.21	-0.63	-4.63	-3.79	-0.59	-3.13	0.00	5.66	0.00
HHSIZE=3	-4.67	-5.17	-2.66	-0.91	-4.86	-2.65	-1.08	-5.85	-0.34	-6.19	-4.56	-7.44	-3.13	-6.99	-4.39	-2.62	-5.79	-2.44	0.27	0.00
HHSIZE=4	0.00	-1.23	-0.80	0.11	-5.93	-5.67	-2.39	-5.34	-1.12	-7.06	-5.53	-6.21	-3.28	-7.14	-5.56	-0.72	-3.29	-0.79	0.79	0.00
HHSIZE $\geq$ 5	0.00	-1.43	-1.01	-0.19	-1.44	-1.87	-1.58	-5.46	-2.41	-5.45	-4.62	-4.61	-1.99	-4.36	-4.02	-1.42	-6.85	-4.67	-0.72	0.00
TEENS=0	-5.89	-4.60	-1.82	2.52	-7.05	-4.42	0.76	-7.43	2.56	-7.47	-4.14	-6.19	-0.34	-5.54	-3.46	0.72	-3.75	-0.33	3.59	0.00
TEENS=1	-2.19	-3.00	-2.17	-1.31	-6.33	-5.74	-4.00	-7.51	-3.53	-8.37	-7.41	-9.64	-6.08	-10.68	-8.26	-4.93	-8.30	-3.59	1.27	0.00
TEENS $\geq$ 2	-0.71	-0.26	-0.40	-0.18	-0.63	-2.53	-2.48	-3.72	-1.48	-5.63	-5.24	-5.59	-3.22	-4.76	-5.19	-1.39	-5.42	-4.03	-0.81	0.00
AGE1=1	0.78	1.52	2.10	3.74	1.73	2.29	3.49	0.00	2.39	-0.80	-0.16	0.24	2.61	-0.58	-0.35	2.61	0.25	1.74	2.21	0.00
AGE2=1	-0.68	-1.43	0.54	2.88	-3.24	-0.61	2.37	-3.30	2.21	-6.19	-4.01	-5.38	-1.20	-4.78	-2.72	0.58	-1.37	-0.07	3.64	0.00
AGE3=1	-7.12	-5.76	-3.61	-0.29	-8.63	-7.71	-3.69	-9.99	-1.26	-9.22	-6.80	-9.66	-4.68	-9.86	-8.20	-4.02	-10.20	-4.99	0.65	0.00
COLLEGE=1	-0.80	0.85	2.22	3.90	-2.38	-0.36	1.98	-3.61	2.37	-2.71	-0.15	-1.89	1.83	-2.55	-1.08	1.69	-2.61	0.48	4.25	0.00
COLLEGE=0	-6.91	-6.63	-4.08	-0.04	-8.47	-7.14	-2.50	-9.47	-0.71	-11.20	-8.96	-11.37	-5.67	-11.05	-9.09	-4.01	-9.28	-4.89	0.69	0.00
MARRIED=1	-6.69	-5.33	-3.35	0.83	-8.33	-6.24	-2.01	-8.81	0.22	-9.86	-7.04	-9.03	-3.46	-9.37	-7.20	-1.43	-7.12	-2.70	2.46	0.00
MARRIED=0	-2.23	-1.92	0.30	2.16	-3.10	-2.15	0.91	-4.85	0.82	-5.00	-3.32	-5.67	-1.70	-5.26	-4.09	-2.40	-6.08	-3.04	1.20	0.00
RETIRED=1	-4.35	-4.77	-3.36	-1.16	-4.74	-4.70	-2.41	-4.62	-0.20	-3.92	-2.70	-3.00	-0.52	-1.98	-1.69	-1.19	-3.13	-1.37	1.73	0.00
RETIRED=0	-4.67	-3.15	-0.73	2.99	-7.14	-4.72	-0.05	-8.91	0.80	-10.31	-7.27	-10.33	-3.92	-10.75	-8.21	-2.22	-8.76	-3.69	2.35	0.00
BLACK=1	-2.16	-4.75	-5.47	-5.04	-6.51	-5.50	-3.54	-5.70	-3.42	-7.63	-6.99	-7.52	-6.38	-7.47	-7.41	-5.91	-7.45	-6.55	-3.44	0.00
BLACK=0	-5.94	-3.96	-0.80	3.53	-7.03	-5.08	-0.23	-8.82	1.54	-9.34	-6.14	-9.04	-2.29	-9.09	-6.53	-0.86	-7.35	-1.91	4.27	0.00
CHURCH=1	-1.92	-3.75	-4.08	-2.27	-5.61	-4.64	-1.83	-4.26	0.91	-4.33	-2.51	-3.25	0.43	-2.27	-2.07	1.20	-4.17	-0.40	2.01	0.00
CHURCH=0	-6.09	-3.92	-0.52	3.40	-6.70	-4.73	-0.32	-9.01	0.27	-10.10	-7.35	-10.32	-4.55	-10.96	-8.28	-3.50	-8.24	-4.23	2.00	0.00
BENEFITS=1	-5.19	-5.38	-3.58	-1.42	-5.14	-5.06	-2.61	-6.06	-1.38	-5.94	-4.72	-5.90	-2.64	-5.02	-4.31	-2.99	-4.63	-2.69	1.47	0.00
BENEFITS=0	-4.03	-2.65	-0.47	3.27	-6.84	-4.36	0.20	-8.07	1.46	-9.29	-6.25	-8.99	-2.99	-9.53	-7.14	-1.36	-8.07	-3.10	2.32	0.00
MOVED=1	-2.38	-1.77	0.43	3.61	-3.92	-1.64	2.18	-5.63	1.94	-6.54	-3.91	-5.48	-0.66	-5.90	-3.87	-0.51	-5.03	-3.24	-0.33	0.00
MOVED=0	-6.27	-5.47	-3.63	-0.66	-7.98	-7.13	-3.60	-8.41	-0.98	-8.91	-6.91	-9.49	-4.66	-9.18	-7.70	-2.89	-7.94	-2.31	4.10	0.00
ONLYMALE=1	1.03	1.69	2.62	3.10	1.39	1.20	2.31	-0.09	2.12	-0.64	0.28	0.00	1.75	0.00	1.07	2.10	0.88	1.84	0.58	0.00
ONLYMALE=0	-6.89	-5.84	-3.08	1.27	-9.06	-6.74	-1.67	-10.20	0.19	-11.07	-7.93	-10.88	-4.27	-10.94	-8.65	-2.91	-9.57	-4.33	2.66	0.00

These ratios are distributed as a studentized maximum modulus distribution [Stoline and Ury (1979)]. With 20 multiple comparisons and infinite degrees of freedom the 5% and 1% approximate critical values are 3.03 and 3.49 respectively.

Table 3b. Tests of First Order Stochastic Dominance. Louisville

Call Range:	0–3	4–5	6–7	8–9	10	11–12	13–14	15	16–19	20	21–22	23–25	26–29	30–31	32–35	35–44	45–50	51–60	61–90	>90
ALL	-8.44	-12.72	-12.64	-8.86	-18.26	-16.14	-12.82	-20.93	-12.60	-22.28	-19.53	-21.49	-14.50	-18.94	-17.38	-11.44	-16.72	-13.54	-5.65	0.00
MARCH	-4.37	-7.28	-7.99	-5.83	-11.48	-9.71	-7.94	-12.30	-7.16	-12.65	-11.35	-11.98	-8.53	-10.87	-10.01	-6.73	-9.97	-8.36	-3.27	0.00
APRIL	-4.74	-7.88	-7.45	-5.22	-10.51	-9.19	-7.40	-11.97	-7.36	-12.92	-11.29	-12.49	-7.94	-10.74	-9.83	-6.42	-9.42	-7.84	-3.56	0.00
MAY	-5.51	-6.88	-6.48	-4.32	-9.67	-9.06	-6.87	-11.97	-7.31	-13.03	-11.19	-12.75	-8.64	-11.18	-10.26	-6.66	-9.57	-7.26	-2.94	0.00
LOW INCOME	-6.15	-8.52	-9.45	-8.21	-13.36	-11.76	-9.61	-14.38	-9.50	-15.10	-13.89	-15.58	-11.41	-13.72	-13.51	-10.32	-13.27	-11.73	-6.45	0.00
HIGH INCOME	-5.81	-9.46	-8.54	-4.60	-12.67	-11.26	-8.72	-15.32	-8.57	-16.50	-13.95	-15.15	-9.54	-13.34	-11.58	-6.54	-10.88	-8.02	-1.92	0.00
HHSIZE=1	-4.58	-6.71	-6.50	-4.21	-8.87	-7.82	-5.64	-9.70	-4.64	-8.57	-6.47	-6.69	-2.99	-5.42	-5.57	-2.34	-5.90	-3.60	-0.65	0.00
HHSIZE=2	-5.39	-7.89	-7.86	-5.25	-11.46	-9.98	-7.93	-12.68	-6.62	-12.98	-11.33	-11.75	-6.88	-10.07	-7.59	-3.37	-5.79	-4.61	-0.79	0.00
HHSIZE=3	-3.70	-5.62	-5.45	-3.48	-8.30	-7.80	-6.72	-11.20	-7.95	-12.15	-11.39	-11.93	-8.14	-8.75	-8.48	-6.07	-7.65	-6.55	-3.28	0.00
HHSIZE=4	-2.69	-4.85	-5.12	-4.98	-7.83	-6.34	-4.93	-7.84	-6.25	-9.89	-8.86	-10.15	-8.05	-9.91	-8.87	-5.79	-8.47	-6.61	-1.63	0.00
HHSIZE $\geq$ 5	-2.55	-4.56	-4.95	-4.69	-6.46	-6.43	-5.92	-7.77	-6.04	-9.47	-8.66	-10.82	-9.51	-11.24	-11.54	-9.98	-11.60	-9.63	-5.52	0.00
TEENS=0	-8.18	-12.21	-12.07	-8.26	-17.34	-15.36	-12.06	-19.88	-11.54	-20.88	-17.91	-18.98	-11.83	-15.91	-13.54	-7.59	-11.58	-8.46	-1.75	0.00
TEENS=1	-2.12	-3.37	-4.15	-3.60	-5.49	-4.98	-3.99	-6.35	-4.92	-7.93	-7.61	-9.84	-7.87	-9.32	-9.96	-7.76	-9.95	-8.57	-5.00	0.00
TEENS $\geq$ 2	-0.58	-1.68	-0.92	-1.37	-3.85	-3.43	-3.60	-4.91	-4.09	-5.72	-5.99	-7.08	-6.86	-8.27	-8.82	-7.58	-10.22	-9.01	-4.63	0.00
AGE1=1	0.64	2.73	1.46	1.56	-2.18	-1.35	0.00	-3.81	-1.50	-4.56	-3.34	-2.58	-0.70	-2.26	-2.03	-2.13	-4.29	-4.35	-1.62	0.00
AGE2=1	-1.83	-4.32	-5.02	-2.69	-7.78	-6.79	-4.50	-8.77	-4.29	-10.87	-8.72	-10.04	-6.13	-9.43	-8.49	-4.72	-8.35	-6.67	-1.61	0.00
AGE3=1	-8.76	-12.81	-12.30	-9.21	-16.66	-14.96	-12.69	-18.92	-12.23	-19.07	-17.41	-19.24	-13.74	-16.58	-15.35	-10.42	-13.86	-10.99	-5.34	0.00
COLLEGE=1	-2.26	-3.01	-2.62	-0.48	-6.84	-4.88	-2.87	-5.51	-1.50	-5.80	-3.82	-4.88	-1.81	-4.60	-3.38	0.06	-2.99	-3.43	-0.69	0.00
COLLEGE=0	-8.60	-12.96	-12.98	-9.88	-17.13	-15.78	-13.08	-20.84	-13.53	-22.19	-20.12	-21.76	-15.46	-18.96	-17.81	-12.92	-17.25	-13.45	-5.89	0.00
MARRIED=1	-4.90	-9.64	-10.27	-7.01	-13.03	-11.58	-9.25	-15.62	-9.98	-17.82	-15.82	-18.48	-13.52	-16.08	-14.22	-9.22	-12.06	-9.71	-3.46	0.00
MARRIED=0	-6.91	-8.81	-8.12	-5.84	-13.08	-11.48	-9.06	-14.21	-8.00	-13.84	-11.88	-11.87	-6.87	-10.60	-10.27	-6.89	-11.62	-9.45	-4.51	0.00
RETIRED=1	-6.12	-8.87	-8.53	-6.59	-11.21	-11.23	-9.94	-13.63	-9.10	-12.55	-11.43	-12.39	-7.85	-9.52	-8.76	-5.15	-5.55	-3.04	-1.65	0.00
RETIRED=0	-5.94	-9.28	-9.49	-6.24	-14.67	-12.13	-9.00	-16.46	-9.42	-18.75	-16.24	-18.03	-12.48	-16.70	-15.33	-10.39	-15.95	-13.40	-5.43	0.00
BLACK=1	-3.72	-7.32	-7.87	-7.49	-10.29	-9.33	-8.15	-10.96	-9.56	-13.93	-12.88	-12.84	-10.54	-12.49	-11.99	-8.82	-11.56	-10.30	-6.21	0.00
BLACK=0	-7.69	-10.94	-10.70	-6.88	-15.87	-13.97	-10.86	-18.44	-10.13	-18.78	-16.22	-18.26	-11.58	-15.53	-13.98	-8.71	-13.18	-10.11	-3.00	0.00
CHURCH=1	-4.17	-5.72	-5.71	-3.62	-6.32	-5.94	-3.85	-6.37	-3.70	-6.36	-5.04	-5.58	-2.42	-4.48	-3.86	-1.44	-3.23	-2.58	0.09	0.00
CHURCH=0	-7.41	-11.39	-11.30	-8.09	-17.17	-15.03	-12.31	-20.07	-12.16	-21.62	-19.20	-21.16	-14.92	-18.87	-17.43	-12.05	-17.05	-13.92	-6.57	0.00
BENEFITS=1	-6.68	-9.11	-9.01	-7.24	-12.04	-11.15	-9.61	-13.92	-9.44	-13.27	-12.07	-12.51	-8.19	-10.80	-10.57	-7.22	-10.51	-8.62	-4.60	0.00
BENEFITS=0	-5.48	-9.04	-9.02	-5.62	-13.87	-11.92	-8.92	-15.91	-8.86	-18.00	-15.50	-17.57	-12.03	-15.66	-13.98	-9.05	-13.27	-10.65	-3.81	0.00
MOVED=1	-2.72	-5.41	-5.39	-2.59	-9.59	-7.49	-5.04	-10.85	-5.47	-13.02	-11.12	-12.89	-8.88	-13.06	-11.87	-7.96	-12.66	-10.00	-4.03	0.00
MOVED=0	-8.45	-11.85	-11.75	-9.11	-15.70	-14.66	-12.37	-18.16	-11.82	-18.16	-16.17	-17.26	-11.49	-13.77	-12.73	-8.25	-11.17	-9.28	-4.00	0.00
ONLYMALE=1	-1.75	-2.20	-1.22	0.66	-3.38	-2.86	-1.85	-4.40	-1.06	-4.60	-3.62	-3.93	-1.14	-2.78	-2.52	-1.71	-4.21	-3.43	-2.52	0.00
ONLYMALE=0	-8.51	-12.97	-13.13	-9.76	-18.29	-16.17	-12.97	-20.66	-12.99	-22.01	-19.41	-21.36	-14.90	-19.02	-17.45	-11.45	-16.22	-13.13	-5.28	0.00

These ratios are distributed as a studentized maximum modulus distribution [Stoline and Ury (1979)]. With 20 multiple comparisons and infinite degrees of freedom the 5% and 1% approximate critical values are 3.03 and 3.49 respectively.

Table 4a. Tests of Second Order Stochastic Dominance. Bowling Green

Call Range:	0–3	4–5	6–7	8–9	10	11–12	13–14	15	16–19	20	21–22	23–25	26–29	30–31	32–35	35–44	45–50	51–60	61–90	>90
ALL	-6.27	-6.42	-5.50	-3.65	-3.80	-5.24	-5.07	-5.29	-5.41	-5.50	-6.33	-7.22	-7.59	-7.74	-8.39	-8.18	-8.16	-8.44	-6.99	0.51
MARCH	-3.02	-3.22	-2.93	-2.09	-2.26	-3.15	-3.12	-3.28	-3.39	-3.44	-3.92	-4.46	-4.70	-4.78	-5.11	-4.94	-4.97	-5.22	-4.51	-0.26
APRIL	-4.56	-4.37	-3.46	-2.16	-2.21	-3.08	-2.97	-3.05	-3.07	-3.13	-3.63	-4.10	-4.28	-4.37	-4.81	-4.80	-4.81	-4.94	-4.06	0.25
MAY	-3.36	-3.57	-3.15	-2.08	-2.11	-2.84	-2.71	-2.83	-2.92	-2.97	-3.43	-3.95	-4.17	-4.25	-4.61	-4.43	-4.36	-4.45	-3.53	0.91
LOW INCOME	-5.00	-5.48	-5.61	-4.90	-5.02	-5.93	-5.93	-6.10	-6.24	-6.27	-6.72	-7.19	-7.31	-7.36	-7.74	-7.65	-7.72	-8.06	-7.37	-1.94
HIGH INCOME	-3.94	-3.73	-2.45	-0.65	-0.76	-1.88	-1.69	-1.84	-1.92	-2.02	-2.74	-3.50	-3.90	-4.05	-4.59	-4.42	-4.37	-4.48	-3.20	2.06
HHSIZE=1	-0.72	-0.73	-0.28	0.51	0.46	-0.24	-0.40	-0.54	-0.94	-0.96	-1.03	-1.11	-1.13	-1.08	-1.06	-0.97	-0.98	-1.12	-1.21	-0.67
HHSIZE=2	-6.28	-5.69	-4.64	-3.11	-3.00	-3.55	-3.11	-3.06	-2.47	-2.41	-2.80	-3.13	-3.15	-3.18	-3.49	-3.45	-3.38	-3.33	-1.99	4.09
HHSIZE=3	-4.67	-5.41	-5.02	-3.96	-3.94	-4.14	-3.71	-3.80	-3.80	-3.84	-4.27	-4.88	-5.27	-5.39	-5.72	-5.53	-5.57	-5.68	-4.82	-0.93
HHSIZE=4	0.00	-0.54	-0.84	-0.67	-1.40	-3.42	-3.82	-3.95	-3.90	-4.03	-4.65	-5.20	-5.35	-5.47	-5.91	-5.46	-5.11	-4.82	-3.61	-0.13
HHSIZE $\geq$ 5	0.00	-0.65	-1.00	-0.88	-0.90	-1.21	-1.43	-1.83	-2.95	-3.15	-3.70	-4.15	-4.14	-4.13	-4.35	-4.09	-4.30	-5.04	-4.83	-1.56
TEENS=0	-5.89	-5.93	-4.96	-3.11	-3.14	-4.19	-3.77	-3.83	-3.60	-3.57	-4.06	-4.47	-4.46	-4.43	-4.65	-4.18	-3.97	-3.94	-2.85	2.14
TEENS=1	-2.19	-2.74	-2.83	-2.54	-3.01	-4.44	-4.86	-5.19	-5.77	-5.93	-6.60	-7.47	-7.99	-8.26	-8.95	-8.85	-8.87	-8.86	-7.03	-0.41
TEENS $\geq$ 2	-0.71	-0.59	-0.51	-0.45	-0.46	-0.90	-1.43	-1.71	-2.19	-2.40	-3.10	-3.84	-4.20	-4.28	-4.65	-4.55	-4.56	-4.97	-4.54	-1.53
AGE1=1	0.78	1.18	1.58	2.18	2.34	2.40	2.68	2.63	2.32	2.22	1.85	1.55	1.59	1.58	1.30	1.34	1.44	1.46	1.80	2.35
AGE2=1	-0.68	-1.09	-0.78	0.22	0.09	-0.57	-0.18	-0.24	-0.36	-0.52	-1.30	-2.08	-2.47	-2.58	-2.90	-2.63	-2.41	-2.26	-1.28	2.70
AGE3=1	-7.12	-7.21	-6.47	-5.03	-5.18	-6.53	-6.64	-6.86	-6.88	-6.86	-7.30	-7.83	-8.08	-8.19	-8.69	-8.61	-8.73	-9.11	-8.00	-1.44
COLLEGE=1	-0.80	-0.20	0.70	1.68	1.52	0.71	0.79	0.59	0.20	0.17	-0.09	-0.28	-0.24	-0.25	-0.54	-0.35	-0.36	-0.54	0.24	3.59
COLLEGE=0	-6.91	-7.46	-6.93	-5.35	-5.42	-6.60	-6.45	-6.58	-6.49	-6.57	-7.40	-8.32	-8.80	-8.96	-9.56	-9.44	-9.43	-9.63	-8.36	-1.48
MARRIED=1	-6.69	-6.80	-6.04	-4.40	-4.56	-5.87	-5.68	-5.82	-5.64	-5.69	-6.39	-7.06	-7.24	-7.33	-7.84	-7.42	-7.24	-7.31	-5.89	0.59
MARRIED=0	-2.23	-2.31	-1.69	-0.66	-0.67	-1.29	-1.18	-1.31	-1.64	-1.71	-2.14	-2.67	-3.02	-3.13	-3.51	-3.71	-3.92	-4.30	-3.80	0.03
RETIRED=1	-4.35	-4.93	-4.86	-4.12	-4.06	-4.52	-4.51	-4.51	-4.11	-4.03	-4.08	-4.06	-3.84	-3.71	-3.62	-3.39	-3.39	-3.46	-2.92	0.57
RETIRED=0	-4.67	-4.49	-3.48	-1.79	-2.02	-3.48	-3.34	-3.61	-4.03	-4.19	-5.12	-6.15	-6.69	-6.92	-7.68	-7.56	-7.55	-7.83	-6.47	0.38
BLACK=1	-2.16	-3.81	-4.94	-5.30	-5.55	-5.95	-5.79	-5.78	-5.73	-5.83	-6.29	-6.76	-7.10	-7.23	-7.51	-7.70	-7.82	-8.09	-7.84	-4.56
BLACK=0	-5.94	-5.68	-4.37	-2.31	-2.39	-3.76	-3.64	-3.88	-4.05	-4.13	-4.87	-5.66	-5.96	-6.07	-6.65	-6.34	-6.26	-6.42	-4.84	2.31
CHURCH=1	-1.92	-2.78	-3.59	-3.62	-3.82	-4.47	-4.35	-4.30	-3.61	-3.51	-3.63	-3.65	-3.33	-3.16	-3.13	-2.58	-2.49	-2.66	-1.85	1.23
CHURCH=0	-6.09	-5.79	-4.33	-2.24	-2.31	-3.60	-3.48	-3.75	-4.25	-4.40	-5.27	-6.25	-6.84	-7.09	-7.84	-7.90	-7.93	-8.15	-6.94	-0.08
BENEFITS=1	-5.19	-5.74	-5.52	-4.65	-4.57	-5.03	-4.98	-5.03	-4.93	-4.91	-5.16	-5.43	-5.49	-5.48	-5.61	-5.52	-5.51	-5.56	-4.74	-0.01
BENEFITS=0	-4.03	-3.85	-2.92	-1.32	-1.56	-3.03	-2.90	-3.14	-3.41	-3.54	-4.38	-5.27	-5.71	-5.90	-6.59	-6.43	-6.43	-6.73	-5.53	0.57
MOVED=1	-2.38	-2.37	-1.62	-0.22	-0.28	-1.11	-0.78	-0.94	-1.33	-1.46	-2.13	-2.76	-2.98	-3.07	-3.52	-3.43	-3.46	-3.82	-3.68	-1.19
MOVED=0	-6.27	-6.50	-5.99	-4.80	-4.95	-6.16	-6.26	-6.40	-6.21	-6.22	-6.72	-7.34	-7.65	-7.76	-8.22	-8.02	-7.96	-8.00	-6.15	1.86
ONLYMALE=1	1.03	1.35	1.85	2.33	2.39	2.22	2.22	2.15	1.91	1.85	1.62	1.42	1.38	1.36	1.29	1.45	1.53	1.59	1.72	1.00
ONLYMALE=0	-6.89	-7.12	-6.27	-4.45	-4.61	-6.02	-5.82	-6.02	-6.06	-6.13	-6.91	-7.76	-8.12	-8.26	-8.90	-8.70	-8.68	-8.97	-7.48	0.36

These ratios are distributed as a studentized maximum modulus distribution [Stoline and Ury (1979)]. With 20 multiple comparisons and infinite degrees of freedom the 5% and 1% approximate critical values are 3.03 and 3.49 respectively.

Table 4b. Test of Second Order Stochastic Dominance. Louisville

Call Range:	0–3	4–5	6–7	8–9	10	11–12	13–14	15	16–19	20	21–22	23–25	26–29	30–31	32–35	35–44	45–50	51–60	61–90	>90
ALL	-8.44	-11.22	-12.94	-12.77	-13.47	-15.57	-16.01	-16.53	-17.75	-18.05	-19.12	-20.26	-20.79	-20.89	-21.33	-20.97	-20.83	-21.04	-19.68	-9.44
MARCH	-4.37	-6.11	-7.44	-7.62	-8.14	-9.48	-9.75	-10.05	-10.65	-10.78	-11.37	-11.96	-12.19	-12.24	-12.47	-12.24	-12.17	-12.33	-11.61	-5.53
APRIL	-4.74	-6.58	-7.70	-7.58	-7.97	-9.11	-9.32	-9.61	-10.30	-10.47	-11.09	-11.74	-12.01	-12.04	-12.27	-12.02	-11.92	-12.03	-11.30	-5.66
MAY	-5.51	-6.74	-7.29	-6.93	-7.24	-8.40	-8.66	-8.97	-9.81	-10.01	-10.68	-11.40	-11.82	-11.90	-12.21	-12.06	-11.99	-12.07	-11.18	-5.15
LOW INCOME	-6.15	-7.81	-9.12	-9.52	-10.12	-11.53	-11.82	-12.11	-12.84	-13.01	-13.65	-14.45	-14.95	-15.06	-15.45	-15.60	-15.68	-16.00	-15.43	-8.95
HIGH INCOME	-5.81	-8.07	-9.23	-8.64	-9.05	-10.64	-10.98	-11.42	-12.44	-12.68	-13.55	-14.39	-14.68	-14.71	-14.98	-14.42	-14.18	-14.18	-12.85	-4.77
HHSIZE=1	-4.58	-6.01	-6.83	-6.65	-6.87	-7.73	-7.86	-8.03	-8.35	-8.36	-8.53	-8.58	-8.34	-8.21	-8.21	-7.92	-7.81	-7.89	-7.32	-2.73
HHSIZE=2	-5.39	-7.07	-8.19	-8.02	-8.50	-9.91	-10.15	-10.45	-10.93	-11.05	-11.65	-12.21	-12.27	-12.23	-12.30	-11.57	-11.17	-10.92	-9.81	-3.37
HHSIZE=3	-3.70	-4.98	-5.69	-5.47	-5.87	-7.12	-7.55	-8.00	-9.29	-9.56	-10.35	-11.18	-11.53	-11.48	-11.46	-11.08	-10.90	-10.83	-10.06	-5.12
HHSIZE=4	-2.69	-3.92	-4.83	-5.20	-5.69	-6.56	-6.61	-6.79	-7.41	-7.63	-8.26	-8.98	-9.50	-9.69	-10.08	-9.90	-9.84	-9.99	-9.07	-3.34
HHSIZE $\geq$ 5	-2.55	-3.64	-4.43	-4.69	-5.02	-5.80	-6.11	-6.35	-7.00	-7.26	-7.98	-8.97	-9.88	-10.22	-10.94	-11.57	-11.90	-12.30	-11.56	-6.78
TEENS=0	-8.18	-10.81	-12.44	-12.23	-12.87	-14.86	-15.27	-15.75	-16.87	-17.12	-18.09	-19.00	-19.25	-19.22	-19.38	-18.58	-18.17	-17.96	-16.21	-5.77
TEENS=1	-2.12	-2.95	-3.65	-3.87	-4.11	-4.66	-4.79	-4.98	-5.61	-5.81	-6.40	-7.31	-8.16	-8.44	-9.11	-9.74	-10.02	-10.51	-10.35	-6.35
TEENS $\geq$ 2	-0.58	-1.15	-1.29	-1.34	-1.77	-2.64	-3.01	-3.27	-3.97	-4.13	-4.61	-5.32	-6.01	-6.33	-7.04	-7.83	-8.35	-9.17	-9.14	-5.55
AGE1=1	0.64	1.50	1.92	1.89	1.46	0.40	0.11	-0.18	-1.19	-1.39	-1.88	-2.20	-2.17	-2.14	-2.24	-2.35	-2.57	-3.07	-3.48	-2.17
AGE2=1	-1.83	-3.15	-4.28	-4.32	-4.72	-5.90	-6.09	-6.32	-6.86	-7.08	-7.84	-8.59	-8.97	-9.08	-9.49	-9.38	-9.35	-9.61	-9.02	-3.56
AGE3=1	-8.76	-11.47	-12.99	-12.81	-13.35	-15.01	-15.42	-15.85	-16.83	-17.02	-17.76	-18.63	-19.09	-19.15	-19.43	-19.04	-18.82	-18.76	-17.32	-8.58
COLLEGE=1	-2.26	-2.77	-2.99	-2.58	-2.93	-4.06	-4.20	-4.31	-4.31	-4.33	-4.55	-4.71	-4.66	-4.63	-4.74	-4.28	-4.04	-4.13	-4.02	-1.63
COLLEGE=0	-8.60	-11.53	-13.34	-13.31	-13.89	-15.63	-16.03	-16.56	-17.95	-18.27	-19.37	-20.56	-21.18	-21.29	-21.71	-21.50	-21.45	-21.62	-20.07	-9.64
MARRIED=1	-4.90	-7.63	-9.66	-9.78	-10.30	-11.77	-11.98	-12.39	-13.46	-13.77	-14.83	-16.08	-16.94	-17.14	-17.56	-17.17	-16.91	-16.83	-15.46	-6.66
MARRIED=0	-6.91	-8.45	-9.13	-8.79	-9.26	-10.77	-11.13	-11.45	-12.09	-12.20	-12.69	-13.06	-12.93	-12.83	-12.99	-12.74	-12.75	-13.08	-12.48	-6.69
RETIRED=1	-6.12	-7.97	-9.03	-8.98	-9.34	-10.57	-11.09	-11.45	-12.23	-12.32	-12.69	-13.12	-13.25	-13.19	-13.22	-12.82	-12.51	-12.12	-10.76	-4.28
RETIRED=0	-5.94	-8.04	-9.46	-9.32	-9.96	-11.75	-11.97	-12.39	-13.47	-13.78	-14.88	-16.04	-16.67	-16.85	-17.43	-17.36	-17.42	-17.95	-17.15	-8.53
BLACK=1	-3.72	-5.86	-7.30	-7.79	-8.27	-9.22	-9.44	-9.67	-10.46	-10.75	-11.54	-12.30	-12.72	-12.88	-13.27	-13.28	-13.31	-13.58	-12.97	-7.89
BLACK=0	-7.69	-9.91	-11.23	-10.87	-11.45	-13.35	-13.73	-14.20	-15.23	-15.44	-16.31	-17.24	-17.63	-17.66	-17.96	-17.50	-17.29	-17.34	-15.93	-6.49
CHURCH=1	-4.17	-5.31	-6.01	-5.80	-5.87	-6.31	-6.21	-6.22	-6.19	-6.18	-6.30	-6.38	-6.17	-6.05	-6.01	-5.51	-5.26	-5.13	-4.29	-0.85
CHURCH=0	-7.41	-9.94	-11.52	-11.41	-12.15	-14.24	-14.77	-15.32	-16.69	-17.01	-18.15	-19.38	-20.07	-20.23	-20.75	-20.59	-20.56	-20.88	-19.87	-10.38
BENEFITS=1	-6.68	-8.44	-9.54	-9.53	-9.94	-11.15	-11.50	-11.82	-12.58	-12.69	-13.10	-13.51	-13.53	-13.48	-13.64	-13.42	-13.35	-13.48	-12.75	-6.92
BENEFITS=0	-5.48	-7.63	-9.00	-8.79	-9.38	-11.14	-11.45	-11.87	-12.89	-13.18	-14.24	-15.37	-16.03	-16.19	-16.65	-16.40	-16.29	-16.47	-15.32	-6.83
MOVED=1	-2.72	-4.22	-5.19	-4.96	-5.43	-6.81	-6.99	-7.31	-8.15	-8.41	-9.32	-10.30	-10.95	-11.20	-11.90	-12.19	-12.41	-12.99	-12.62	-6.43
MOVED=0	-8.45	-10.81	-12.28	-12.24	-12.78	-14.39	-14.82	-15.24	-16.17	-16.35	-17.02	-17.71	-17.90	-17.83	-17.86	-17.17	-16.80	-16.59	-15.12	-6.94
ONLYMALE=1	-1.75	-2.13	-2.05	-1.48	-1.53	-2.09	-2.24	-2.39	-2.66	-2.72	-3.01	-3.27	-3.29	-3.24	-3.29	-3.23	-3.32	-3.63	-3.95	-3.36
ONLYMALE=0	-8.51	-11.38	-13.29	-13.30	-14.02	-16.01	-16.39	-16.87	-18.03	-18.31	-19.33	-20.43	-20.96	-21.07	-21.50	-21.10	-20.88	-20.96	-19.38	-8.98

These ratios are distributed as a studentized maximum modulus distribution [Stoline and Ury (1979)]. With 20 multiple comparisons and infinite degrees of freedom the 5% and 1% approximate critical values are 3.03 and 3.49 respectively.

Table 5. Simulation Results

BOWLING GREEN					
Tariff		$\varepsilon = -0.10$	$\varepsilon = -0.17$	$\varepsilon = -0.45$	$\varepsilon = -0.70$
Ex–Post	A	74.911	44.065	16.647	10.702
	p	0.119	0.070	0.026	0.017
	x	134.340	134.340	134.340	134.340
	V	19.127	11.251	4.250	2.732
	π	76.357	44.916	16.968	10.908
	W	95.484	56.167	21.219	13.641
Op. TPT	A	77.047	45.322	17.122	11.007
	p	0.079	0.046	0.018	0.011
	x	136.312	136.312	136.312	136.312
	V	18.371	10.807	4.083	2.625
	π	79.389	46.699	17.642	11.341
	W	97.760	57.506	21.724	13.966
Op. NLT	A	77.134	45.373	17.141	11.019
	p	0.141	0.083	0.031	0.020
	x	141.431	141.431	141.431	141.431
	V	21.868	12.863	4.859	3.124
	π	68.936	40.551	15.319	9.848
	W	90.803	53.414	20.179	12.972
LOUISVILLE					
Tariff		$\varepsilon = -0.10$	$\varepsilon = -0.17$	$\varepsilon = -0.45$	$\varepsilon = -0.70$
Ex–Post	A	100.530	59.135	22.340	14.361
	p	0.195	0.115	0.043	0.028
	x	174.076	174.076	174.076	174.076
	V	21.323	12.543	4.739	3.046
	π	103.002	60.590	22.889	14.715
	W	124.326	73.133	27.628	17.761
Op. TPT	A	108.266	63.686	24.059	15.467
	p	0.119	0.070	0.026	0.017
	x	179.017	179.017	179.017	179.017
	V	17.046	10.027	3.788	2.435
	π	110.498	64.999	24.555	15.785
	W	127.543	75.026	28.343	18.221
Op. NLT	A	115.539	67.964	25.675	16.506
	p	0.119	0.070	0.026	0.017
	x	179.017	179.017	179.017	179.017
	V	9.773	5.749	2.172	1.396
	π	114.385	67.286	25.419	16.341
	W	124.158	73.034	27.591	17.737

Average value of 10,000 random draws from kernel estimates of the corresponding probability density functions.

Figure 1. Empirical Distributions

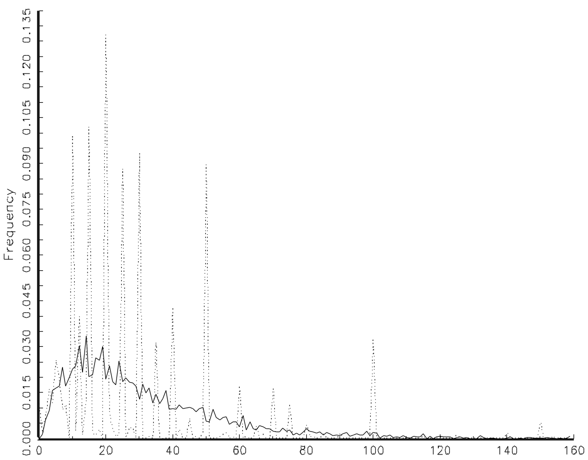

FIGURE 1.1: Actual and Expected (dotted) Calls in BG

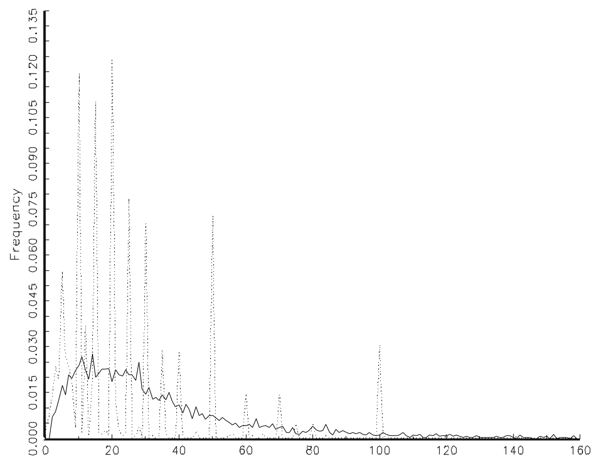

FIGURE 1.2: Actual and Expected (dotted) Calls in LV

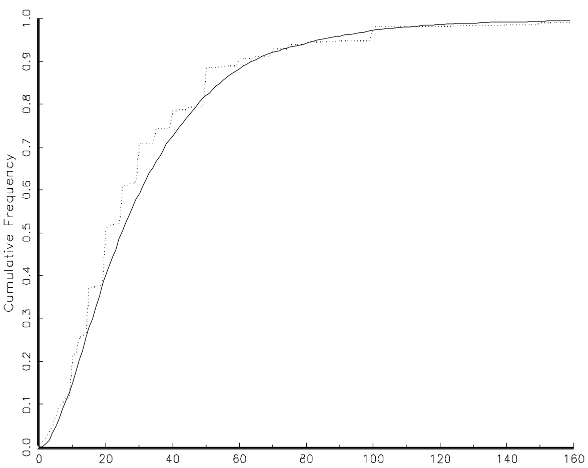

FIGURE 1.3: Actual and Expected (dotted) Calls in BG

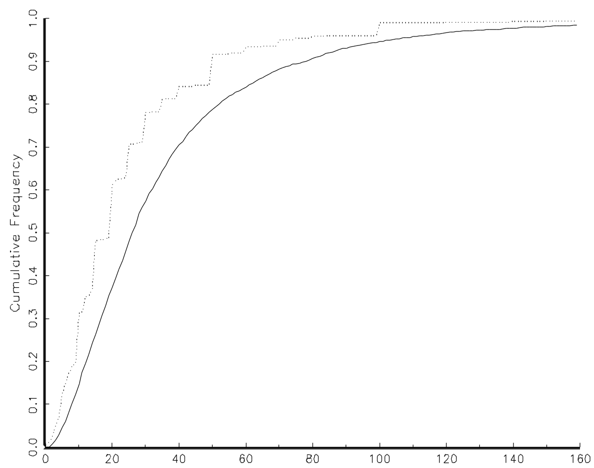

FIGURE 1.4: Actual and Expected (dotted) Calls in LV

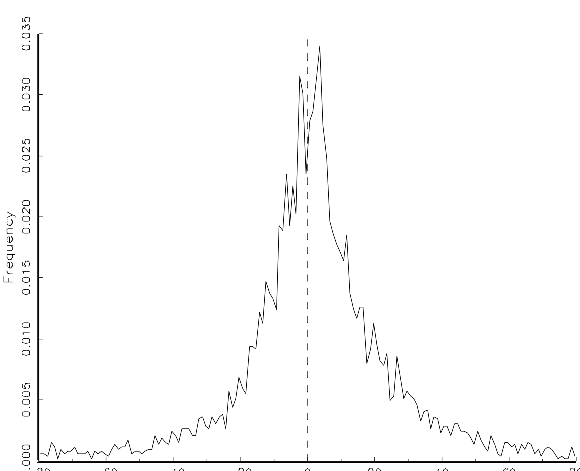

FIGURE 1.5: Call Expectation Bias in BG

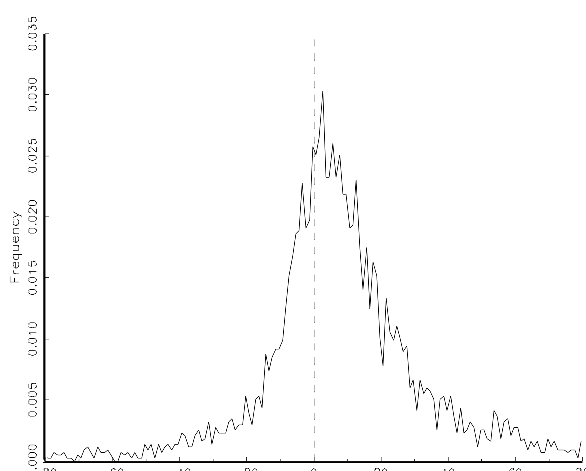

FIGURE 1.6: Call Expectation Bias in LV

Figure 2. Hazard Rate Dominance

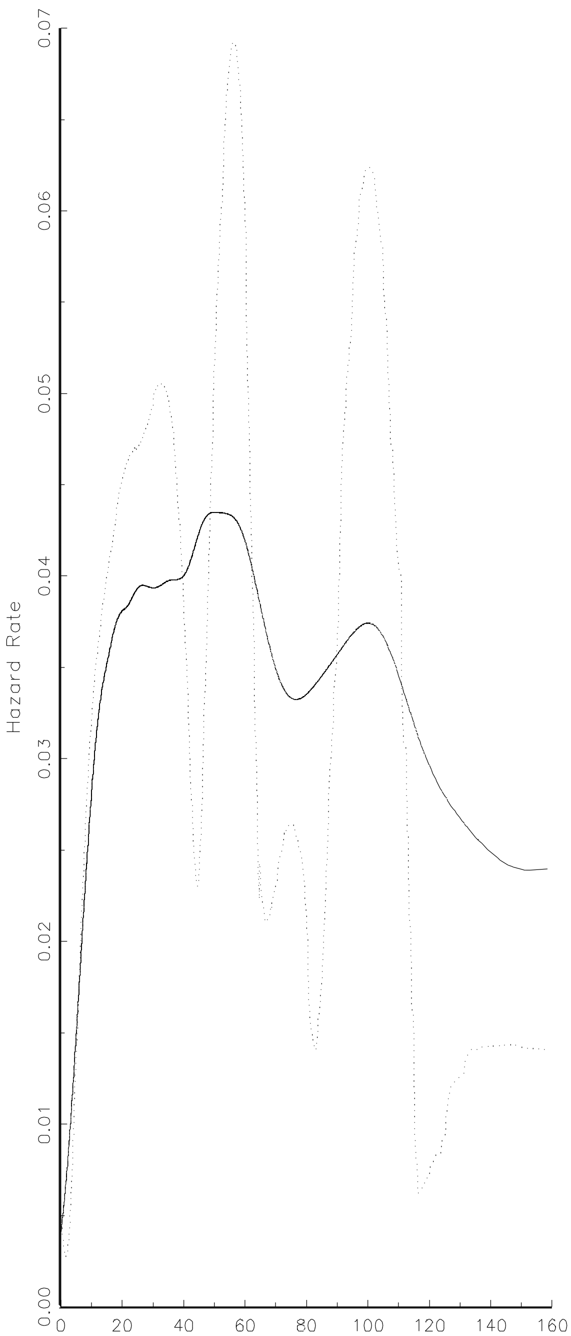

FIGURE 2.1: Actual and Expected (dotted) Calls in BG

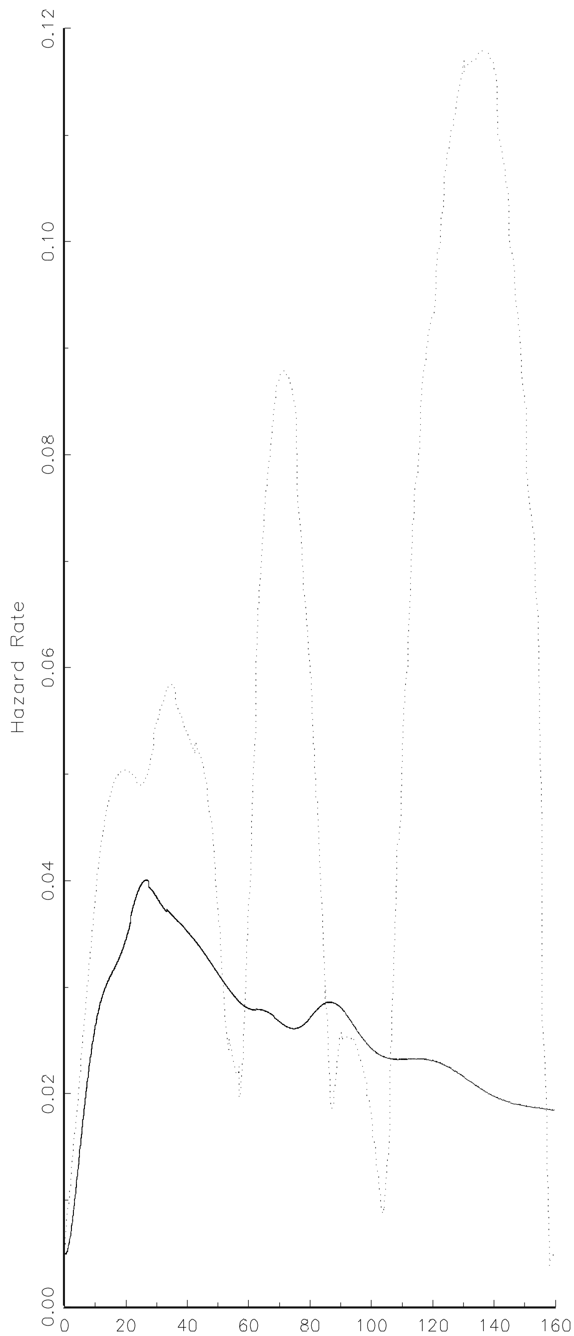

FIGURE 2.2: Actual and Expected (dotted) Calls in LV

Figure 3. Bowling Green: Kernel Estimates

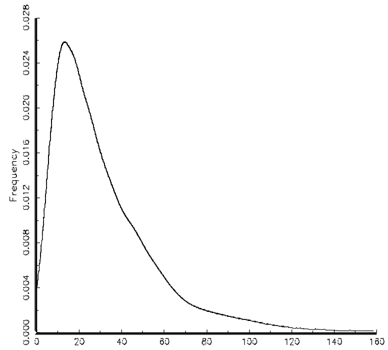

FIGURE 3.1: Actual Number of Calls

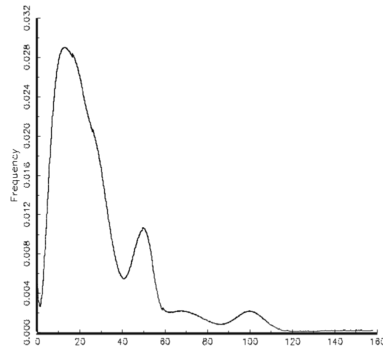

FIGURE 3.2: Expected Number of Calls

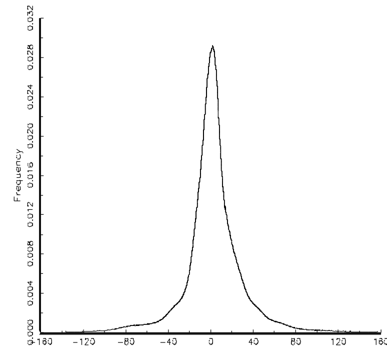

FIGURE 3.3: Call Expectation Bias

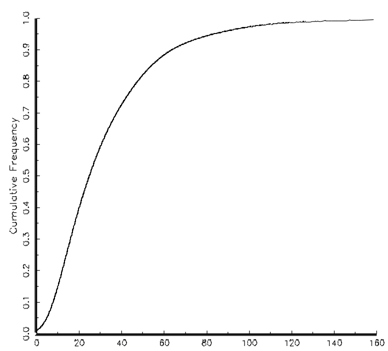

FIGURE 3.4: Actual Number of Calls

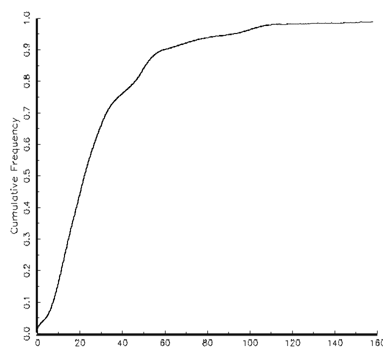

FIGURE 3.5: Expected Number of Calls

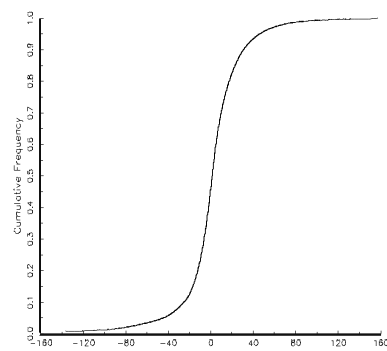

FIGURE 3.6: Call Expectation Bias

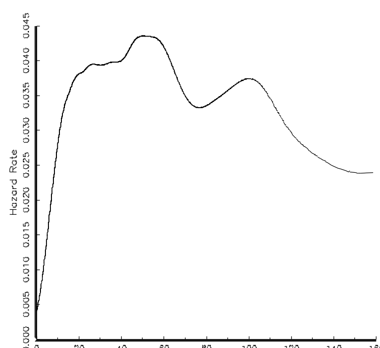

FIGURE 3.7: Actual Number of Calls

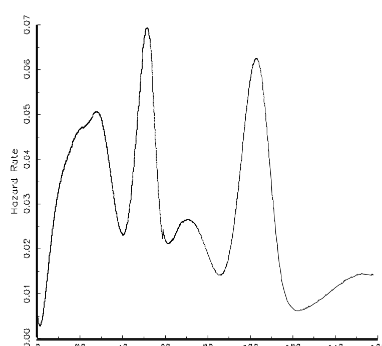

FIGURE 3.8: Expected Number of Calls

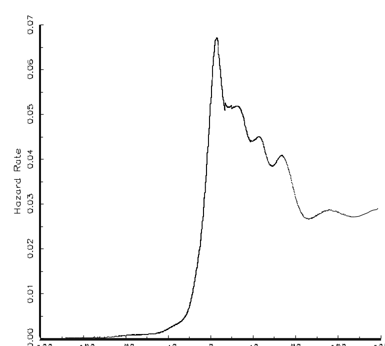

FIGURE 3.9: Call Expectation Bias

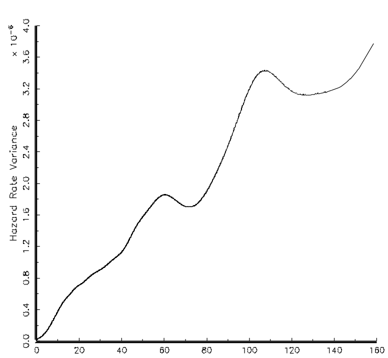

FIGURE 3.10: Actual Number of Calls

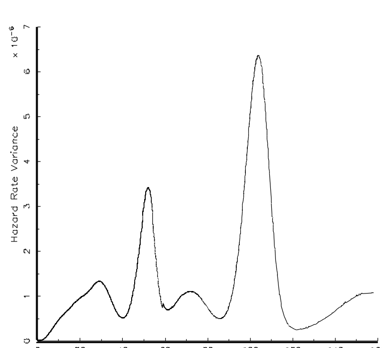

FIGURE 3.11: Expected Number of Calls

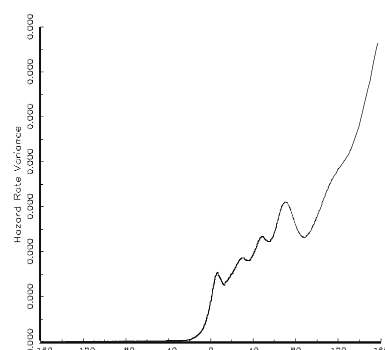

FIGURE 3.2: Call Expectation Bias

Figure 4. Louisville: Kernel Estimates

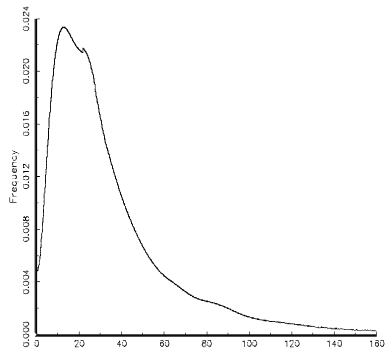

FIGURE 4.1: Actual Number of Calls

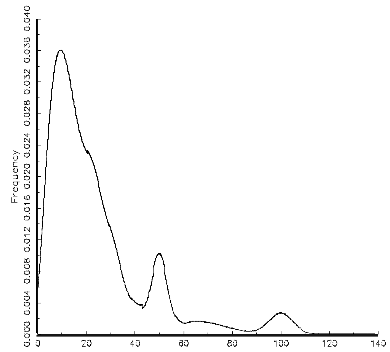

FIGURE 4.2: Expected Number of Calls

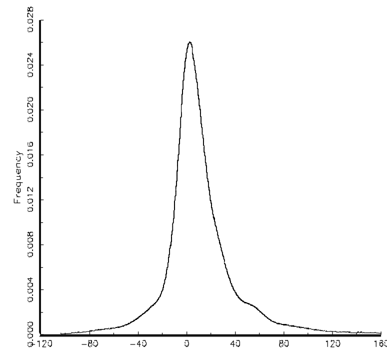

FIGURE 4.3: Call Expectation Bias

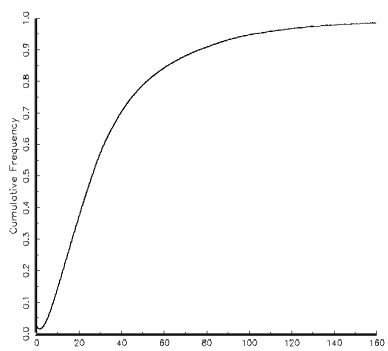

FIGURE 4.4: Actual Number of Calls

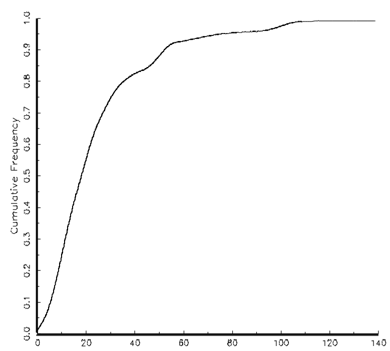

FIGURE 4.5: Expected Number of Calls

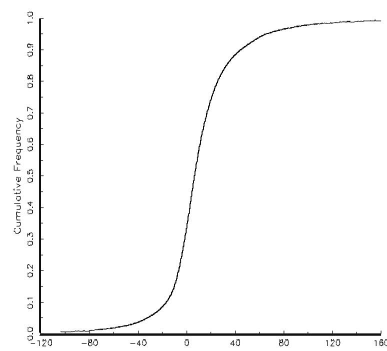

FIGURE 4.6: Call Expectation Bias

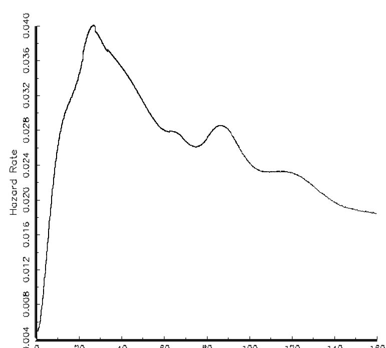

FIGURE 4.7: Actual Number of Calls

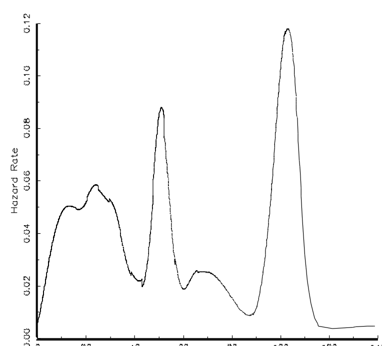

FIGURE 4.8: Expected Number of Calls

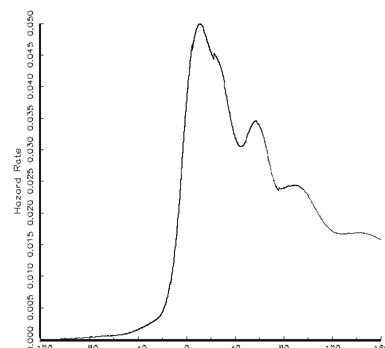

FIGURE 4.9: Call Expectation Bias

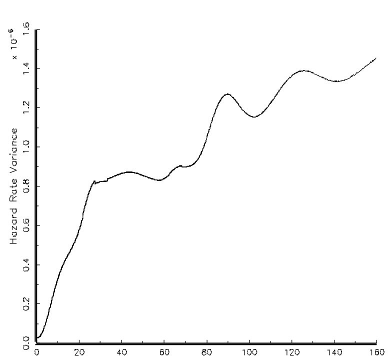

FIGURE 4.10: Actual Number of Calls

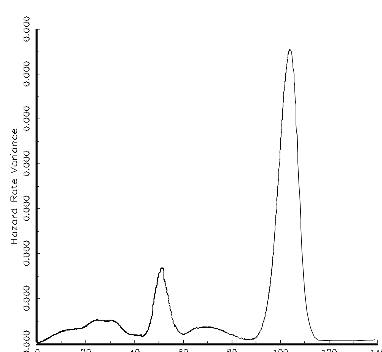

FIGURE 4.11: Expected Number of Calls

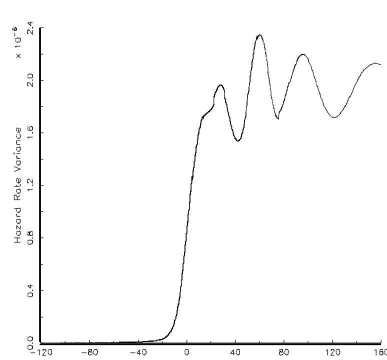

FIGURE 4.12: Call Expectation Bias