GOVERNANCE AND THE EFFICIENCY OF ECONOMIC SYSTEMS
GESY

Discussion Paper No. 330

Uncertain Demand, Consumer Loss Aversion, and Flat-Rate Tariffs

*Fabian Herweg

* University of Bonn

July 2010

Financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15 is gratefully acknowledged.

Sonderforschungsbereich/Transregio 15 · www.sfbtr15.de
Universität Mannheim · Freie Universität Berlin · Humboldt-Universität zu Berlin · Ludwig-Maximilians-Universität München
Rheinische Friedrich-Wilhelms-Universität Bonn · Zentrum für Europäische Wirtschaftsforschung Mannheim

Speaker: Prof. Dr. Urs Schweizer. · Department of Economics · University of Bonn · D-53113 Bonn,
Phone: +49(0228)739220 · Fax: +49(0228)739221

Uncertain Demand, Consumer Loss Aversion, and Flat-Rate Tariffs*

FABIAN HERWEG†

July 16, 2010

The so called flat-rate bias is a well documented phenomenon caused by consumers' desire to be insured against fluctuations in their billing amounts. This paper shows that expectation-based loss aversion provides a formal explanation for this bias. We solve for the optimal two-part tariff when contracting with loss-averse consumers who are uncertain about their demand. The optimal tariff is a flat rate if marginal cost of production is low compared to a consumer's degree of loss aversion and if there is enough variation in the consumer's demand. Moreover, if consumers differ with respect to the degree of loss aversion, firms' optimal menu of tariffs typically comprises a flat-rate contract.

JEL classification: D11; D43; L11

Keywords: Consumer Loss Aversion; Flat-Rate Tariffs; Nonlinear Pricing; Uncertain Demand

1. INTRODUCTION

Nowadays, flat-rate tariffs are widely utilized in many industries, e.g., telephone services, Internet access, car rental, car leasing, DVD rental, amusement parks, health clubs, and many others. With a flat-rate tariff a consumer pays a fixed amount, which is independent of his usage, to obtain unlimited access to a good or a service. The fact that flat-rate tariffs are such favorable pricing schemes is hard to reconcile with orthodox economic theory, in particular for industries where marginal costs are non-negligible. If marginal costs are positive, a marginal payment of zero leads to an inefficiently high level of consumption which hardly can be optimal. On the other hand, usage-based pricing may cause positive transaction costs for measuring the actual usage of a consumer.[1] In many of the examples provided above, however, marginal costs of production or service provision are positive but transaction costs for measuring usage are close to zero.[2] This holds true, for instance, for amusement parks the leading example of flat-rate pricing in the IO literature. The usage

*In preparing this paper I have greatly benefited from comments made by Paul Heidhues, David Laibson, Kristóf Madarász, Daniel Müller, Martin Peitz, Matthew Rabin,Klaus M. Schmidt, Urs Schweizer, Igor Sloev, and Philipp Weinschenk. I also thank participants at the Nordic Conference on Behavioral and Experimental Economics at Oslo (2009), at the IMEBE at Bilbao (2010), at the SFB Meeting at Caputh (2010), and at the WZB Conference Biases in Markets, Berlin (2010). All remaining errors are my own responsibility. No-one else can be blamed for any shortcomings.

†University of Bonn, Department of Economics, Chair of Economic Theory II, Adenauerallee 24-42, D-53113 Bonn, Germany, E-mail address: fherweg@uni-bonn.de.

[1] Sundararajan (2004) shows that it is always optimal for a monopolistic firm to offer a flat rate next to usage-based tariffs if marginal costs are zero and there are positive transaction costs accompanied with usage-based pricing.

[2] It is important to notice that, despite conventional wisdom, the marginal cost of a telephone call is not zero, see for instance Faulhaber and Hogendorn (2000). Moreover, one should keep in mind that telephone companies pay access charges on a per minute basis for off-net calls.

of flat-rate contracts presumably is most puzzling for rental cars: The price for a rental car typically is fixed per day and does not depend on the mileage.[3] The costs for the rental car company are clearly higher if the car is used more heavily, due to for instance a higher wear of the tires. To ascertain how many miles a customer drove with the car is relatively easy and not very costly for the company.[4] Given these observations, what reasons do firms have to offer flat-rate contracts? This paper provides a theoretical answer to that question, which lies outside standard consumer behavior.[5]

There is plenty of evidence, that consumers facing the choice between several tariffs often do not select the optimal one given their consumption patterns. In particular, consumers often prefer a flat-rate tariff even though they would save money with a measured tariff. Train (1991) referred to this phenomenon as "flat-rate bias".[6] Given the fact that consumers are willing to pay a "flat-rate premium" it is unsurprising that this tariff form is widely utilized in many industries. The question is, however, what causes the flat-rate bias? Train et al. (1989) point out that "customers do not choose tariffs with complete knowledge of their demand, but rather choose tariffs [...] on the basis of the insurance provided by the tariff in the face of uncertain consumption patterns". Standard risk aversion, however, cannot capture this insurance motive, since the variations in billing rates are usually small compared to a consumer's income.[7] Already Train (1991) states that "[t]he existence of this [flat-rate] bias is problematical. Standard theory of consumer behavior does not incorporate it." Therefore, to capture first-order risk aversion we posit that consumers are loss averse.[8] A loss-averse consumer dislikes even small deviations from his reference point. In our model, the consumer's demand is uncertain at the point where he selects a tariff. We assume that the consumer forms rational expectations about his invoice, which determine his reference point.[9] The consumer feels a loss if his actual invoice amount is above his reference point, and he feels a gain if it is below his reference point. We follow Kőszegi and Rabin (2006, 2007) and assume that the reference point is a full distribution of the possible billing rates. To illustrate this concept, suppose the consumer's amount invoiced is either \$15, \$20, or \$30. Then, receiving a bill of \$20 generates a mixture of feelings, a gain of \$5 and a loss of \$10. We show that a consumer with these preferences is biased in favor of flat-rate tariffs, since flat rates insure against the risk of losses in periods of greater than average consumption.

First, we consider a monopolist who offers a two-part tariff to an ex ante homogeneous group of consumers. After accepting the contract, each consumer privately observes his

[3] For instance, in Germany, the rental car companies Sixt, Europcar, and Hertz (the three major brands) offer flat-rate contracts. Another common contractual form for rental cars is a three-part tariff: the contract includes a mileage allowance and a charge per mile thereafter.

[4] Other well-fitting examples are the flat-rate contracts for leasing cars newly introduced by Ford and Citroën in Germany. These contracts cover—next to the usual services—also wear repairings for a fixed amount per month that does not depend on the mileage.

[5] We do not claim that our explanation is the only explanation for observing flat-rate tariffs. Another explanation may be preferences for larger choice sets or mental accounting, see Thaler (1999).

[6] For an overview of the empirical evidence documenting the flat-rate bias see Lambrecht and Skiera (2006).

[7] Cf. Clay et al. (1992) or Miravete (2002).

[8] That consumer loss aversion could potentially explain the flat-rate bias is argued in the marketing literature, see Lambrecht and Skiera (2006).

[9] Evidence for the assumption that expectations determine the reference point is documented by Abeler et al. (forthcoming), Post et al. (2008), Crawford and Meng (2009), Gill and Prowse (2009), and Ericson and Fuster (2009).

demand type and thereafter chooses a quantity. The main finding is that the optimal tariff is a flat rate if (i) the consumers are sufficiently loss averse, (ii) consumers' consumption patterns are uncertain, and (iii) the firm's marginal cost of production or service provision is low.

We extend the model to an imperfectly competitive market where firms compete for loss-averse consumers. Moreover, we allow for consumer heterogeneity with respect to their degree of loss aversion. Now, each firm offers a menu of two-part tariffs to screen differently loss-averse consumers. The symmetric information case in which firms observe a consumer's degree of loss aversion, as well as the asymmetric information case in which the degree of loss aversion is private information, is analyzed. In equilibrium of the symmetric information benchmark, firms offer a flat-rate tariff to those consumers whose degree of loss aversion compared to marginal cost exceeds a certain threshold. Consumers with a lower degree of loss aversion are assigned to a metered tariff, i.e., a two-part tariff with a strictly positive unit price. These findings turn out to carry over likewise to the asymmetric information case. Moreover, under certain conditions, differently loss averse consumers do not impose informational externalities on each other. In other words, firms may be able to screen a consumer's degree of loss aversion at no cost. This finding has the interesting implication that a firm's optimal menu of two-part tariffs comprises a flat-rate contract if demand is uncertain and at least some consumers are sufficiently loss averse.

The paper proceeds as follows. Section 2 presents a simple example that illustrates the main findings. Section 3 presents the baseline model with one firm and homogeneous consumers. The demand function of a loss averse consumer for a given tariff is analyzed in Section 4. Section 5 identifies the conditions under which a flat-rate tariff is optimal. Section 6 extends the model to allow for imperfect competition and heterogeneous consumers. The literature documenting evidence for the flat-rate bias as well as related theoretical articles are reviewed in Section 7. Section 8 concludes.

2. ILLUSTRATIVE EXAMPLE

In the following, we present a simple example that illustrates our main findings. Consider a monopolist who sells one good to a single consumer. The monopolist produces with constant per-unit cost $0 < c < 6$. Suppose the take-it-or-leave-it offer the monopolist makes to the consumer is a two-part tariff (p, L), where p is the price per unit and L is the basic charge. For simplicity assume that the offered tariff is either cost based ($p = c$) or a flat rate ($p = 0$). The consumer's demand—his satiation point—depends on his demand type $\theta = 6, 10$. With probability α the type is "low demand" (L) and $\theta = 6$, while with probability $1 - \alpha$ the type is "high demand" (H) and $\theta = 10$. The consumer's (intrinsic) utility is quasi linear and given by $\theta q - (1/2)q^2 - pq - L$, where q denotes the quantity. If the demand type is H, then the quantity the consumer demands under the cost-based and the flat-rate tariff is $q_H(c) = 10 - c$ and $q_H(0) = 10$, respectively. The corresponding gross benefit of consuming the good is $u_H(c) = (1/2)(100 - c^2)$ under the cost-based tariff and $u_H(0) = (1/2)100$ under a flat rate. If the type is L then the demanded quantities and the corresponding gross benefits under the cost-based and the flat-rate contract are: $q_L(c) = 6 - c$, $u_L(c) = (1/2)(36 - c^2)$ and $q_L(0) = 6$, $u_L(0) = (1/2)36$, respectively.

At date 1, when the consumer decides whether or not to accept the monopolist's offer, neither the consumer nor the firm knows the demand type. If the consumer rejects the offer, then he receives his reservation utility of zero. At the beginning of date 2, the consumer learns his type and thereafter he makes his quantity choice. Let $S_i(p) := u_i(p) - cq_i(p)$ with $i = H, L$ be the generated surplus for the per-unit price p and a given demand type of the consumer. If the consumer has "standard preferences", it is optimal for the monopolist to offer the cost-based tariff. Via the lump sum fee, the monopolist can extract the whole (expected) generated surplus. A cost-based tariff provides the correct incentives for the consumer to choose for each realized type the efficient quantity. With a flat rate, demanded quantities are too high since marginal prices do not reflect production costs. The difference in profits from the cost-based tariff to the flat rate is

$$\pi^{USAGE} - \pi^{FLAT} = \alpha[S_L(c) - S_L(0)] + (1-\alpha)[S_H(c) - S_H(0)] = (1/2)c^2.$$

Now, suppose the consumer is expectation-based loss averse with respect to his billing amount according to Kőszegi and Rabin (2006). Next to standard (intrinsic) utility from consuming the good and paying the bill, the consumer derives gain-loss utility by comparing his actual billing amount with his lagged rational expectations. More precisely, the consumer evaluates a given amount of the bill by comparing it to all possible billing amounts, where each comparison is weighted with the probability with which the alternative outcome occurs ex ante. Moreover, let $\lambda > 1$ be the weight the consumer attaches to losses, while the weight on gains equals 1. To keep the example as simple as possible, it is assumed that the consumer ignores his gain-loss utility when making his quantity choice. Thus, the demanded quantities under both tariffs and for each demand type are the same as before. When deciding whether or not to accept the contract, the consumer takes his expected gain-loss utility into account.[10]

First, consider the case where the consumer signed the measured tariff option. If the realized type is H then the consumer's utility amounts to

$$\underbrace{u_H(c) - cq_H(c) - L_c}_{\text{intrinsic utility}} - \underbrace{\alpha\lambda[cq_H(c) - cq_L(c)]}_{\text{gain-loss utility}}.$$

The first term is the gross utility from consuming amount $q_H(c)$ in the high-demand state minus the tariff payment to the firm. The second term represents the perceived loss. With probability α the consumer expected that his type is low demand which corresponds to a billing amount of $L_c + cq_L(c)$. The actual type, however, is high demand and thus his billing amount is $L_c + cq_H(c)$. Comparing the expected bill with the actual bill leads to a loss of $cq_H(c) - cq_L(c)$ which is weighted with λ. With probability $1-\alpha$ the consumer expected that his type is high demand. Since the actual type is H this comparison does neither lead to the sensation of a gain nor of a loss.

The consumer's utility if demand type L is realized amounts to

$$\underbrace{u_L(c) - cq_L(c) - L_c}_{\text{intrinsic utility}} + \underbrace{(1-\alpha)[cq_H(c) - cq_L(c)]}_{\text{gain-loss utility}}.$$

[10]In this example, we do not solve for a personal equilibrium. Nevertheless, the consumer forms rational expectations about his future consumption.

The first term is the standard consumption utility from consuming the low quantity and making the corresponding payment. The second term captures the consumer's perceived gain. With probability $1-\alpha$ the consumer expected to purchase the high quantity and paying a high bill, which feels like a gain when comparing it to the actual billing amount. With probability α the consumer expected to pay exactly his actual amount of the bill. In this case, this comparison does neither lead to the sensation of a loss nor of a gain.

The consumer is fully rational and in consequence he anticipates ex ante these comparisons of billing amounts. Thus, the consumer's expected utility from accepting the measured tariff is

$$\alpha S_L(c) + (1-\alpha)S_H(c) - L_c - (\lambda - 1)\alpha(1-\alpha)4c.$$

The consumer expects to incur a net loss which reduces his utility, since losses loom larger than gains of equal size. Put verbally, a loss averse consumer dislikes fluctuations in his billing amount. A flat rate, in contrast, completely insures the consumer and he does not expect to incur any gain-loss sensation. The consumer's expected utility from signing the flat-rate tariff is

$$\alpha u_L(0) + (1-\alpha)u_H(0) - L_0.$$

The term $(\lambda-1)\alpha(1-\alpha)4c$ captures the "flat-rate premium" the consumer is willing to pay more for the flat rate than for the usage-based tariff in addition to his increased willingness to pay due to higher consumption.

The profits generated by the flat-rate tariff exceed the profits of the cost-based tariff if

$$(\lambda - 1)\alpha(1-\alpha)8 > c.$$

Hence, a flat rate is optimal when three criteria are satisfied: (i) the consumer is loss averse, i.e., $\lambda > 1$, (ii) the per-unit production costs are not too high, and (iii) the consumption pattern is sufficiently uncertain, i.e., α is neither close to 0 nor to 1. The model predicts, for instance, that one observes flat-rate contracts for rental cars, in particular at vacation resorts where customers are unfamiliar with the network of roads. The model does not predict flat rates for heating oil. Typically, the demand for heating oil is uncertain but the marginal costs are high.

A further interesting insight can be obtained by considering the case where the monopolist sells to two ex ante heterogeneous groups of consumers. Consumers in the first group have standard preferences while consumers belonging to the other group are loss averse. Suppose that under complete information it is optimal to offer a flat-rate tariff to the group of loss averse consumers. What is the optimal menu of two-part tariffs if the consumers are privately informed about their degree of loss aversion? The monopolist optimally offers the cost-based and the flat-rate tariff as if it could observe consumers' types. By doing so, the loss-averse consumers strictly prefer the flat rate. The standard types weakly prefer the cost-based tariff, their expected utility equals zero under both contracts.[11] Thus, the monopolist can screen differently loss-averse consumers at no cost. The reason is that the expected utility from signing a flat-rate contract is independent of a consumer's degree of loss aversion.

[11] In a perfectly competitive market not only loss averse consumers strictly prefer the flat-rate but also the standard consumers strictly prefer the cost-based tariff. In a perfectly competitive market the basic charge is determined by a zero profit condition.

3. Monopolistic Market with Homogeneous Consumers

3.1. Players and Timing

We consider a market where a monopolist produces a single good at constant marginal cost $c > 0$ and without fixed costs. The monopolist offers a two-part tariff to a continuum of ex ante homogeneous consumers of measure one. The tariff is given by $T(q) = L + pq$, where $q \geq 0$ is the quantity and L and p denote the basic charge and the per-unit price, respectively. At the contracting stage a consumer does not know his future demand type $\theta \in [\underline{\theta}, \bar{\theta}] \equiv \Theta$. Consumers' demand types are independently and identically distributed according to the commonly known and twice differentiable cumulative distribution function $F(\cdot)$.[12] Let the probability density function be $f(\cdot)$. To make this assumption more vivid consider a consumer who decides today whether or not to sign a contract with a car rental company for his holidays in a few weeks. How frequently he will use the rented car depends on the weather. If the sun is always shining the consumer uses the car only to drive to the nearby beach. But if the weather is bad he takes longer sight-seeing trips.

The sequence of events is as follows: (1) The monopolist makes a take-it-or-leave-it offer (L, p) to consumers. (2) Each consumer forms expectations about his demand and decides whether or not to accept the offered two-part tariff. (3) At the beginning of stage 3, each consumer privately observes his demand type θ. Thereafter, each consumer, who accepted the offer, demands a quantity that maximizes his utility. (4) Finally, payments are made according to the demanded quantities and the concluded contracts.

3.2. Consumers' Preferences

We assume that consumers are loss averse, in the sense that a consumer is disappointed if the payment he has to make exceeds his reference payment. For instance, consumers typically feel a loss if at the end of the month the invoice from their telecommunication provider is larger than expected. Since, for the situations we have in mind, it is natural to assume that the reference point incorporates forward looking expectations, we apply the approach of reference-dependent preferences developed by Kőszegi and Rabin (2006, 2007). First, this concept posits that overall utility has two additively separable components, consumption utility (intrinsic utility) and gain-loss utility. Second, the consumer's reference point is determined by his rational expectations about outcomes. Finally, a given outcome is evaluated by comparing it to each possible outcome, where each comparison is weighted with the ex-ante probability with which the alternative outcome occurs.

The consumer's intrinsic utility is quasi linear in money; formally, intrinsic utility equals $u(q, \theta) - T(q)$ if he accepts the contract.[13]

For the markets we have in mind, like rental cars or Internet services, even if the price per unit is zero, demand is bounded. Therefore, we assume that there exists a satiation

[12] All our findings are robust to assuming that consumers' demand types are perfectly correlated, i.e., θ is rather a state of the world than a demand type.

[13] With the utility function being quasi linear, the consumer is not risk averse in the usual sense. Standard risk aversion cannot explain flat-rate contracts, since a risk- but not loss-averse consumer is locally risk neutral. Hence, a marginal price slightly above zero creates incentives to reduce overconsumption without reducing the consumer's expected utility. We focus on pure loss aversion to highlight the effect of loss aversion on the optimal pricing scheme.

point, $q^S(\theta)$, and that overconsumption is harmless, i.e., free disposal is possible. Additionally, it is assumed that a higher demand type is associated with a stronger need for the good. The assumptions imposed on the consumer's intrinsic utility function concerning the consumption good are summarized as follows:

Assumption (A1) *For all $\theta \in \Theta$, (i) not consuming the good yields zero intrinsic utility $u(0,\theta) \equiv 0$, and (ii) $u(q,\theta)$ is C^3 for $q \leq q^S(\theta)$. Furthermore, intrinsic utility has the following properties,*

$$\begin{aligned} &\partial u(q,\theta)/\partial q > 0 \quad \text{for } q < q^S(\theta), && \partial^2 u(q,\theta)/\partial q^2 < 0 \quad \text{for } q \leq q^S(\theta), \\ &\partial u(q,\theta)/\partial q = 0 \quad \text{for } q \geq q^S(\theta), && \partial^2 u(q,\theta)/\partial q \partial\theta > \kappa \quad \text{for } q \leq q^S(\theta), \end{aligned}$$

with $\kappa > 0$. (iii) $\partial u(0,\theta)/\partial q = \infty$.

Note that $\partial u(q,\theta)/\partial\theta > 0$ for $q \in (0, q^S(\theta)]$, since the intrinsic utility of zero consumption is normalized to zero and marginal utility is increasing in the demand type. By Assumption (A1), the satiation point $q^S(\theta)$ is defined by

$$q^S(\theta) = \min\{q \in \mathbb{R}^+ | \partial u(q,\theta)/\partial q = 0\} \ . \tag{1}$$

The satiation point is increasing in the demand type θ, formally:[14]

$$\frac{dq^S(\theta)}{d\theta} = -\frac{\partial^2 u(q^S(\theta),\theta)/\partial q \partial\theta}{\partial^2 u(q^S(\theta),\theta)/\partial q^2} > 0 \ . \tag{2}$$

As a special case of (A1), which we will make use of in a later section to illustrate our findings, consider $u(q,\theta) = \theta q - (1/2)q^2$ for $q \leq \theta = q^S(\theta)$ and $u(q,\theta) = (1/2)\theta^2$ for $q > \theta$. The Inada condition (iii) is replaced by $\underline{\theta} > \varepsilon > 0$, which guarantees that the consumer demands a positive quantity for each realization of the demand type.

For simplicity, we depart from the Kőszegi and Rabin concept by assuming that the consumer feels gains and losses only in the money dimension.[15] We discuss the implications of this assumption further below. Put verbally, we posit that a consumer does not feel a loss if the weather is nice and he uses the rental car less often than expected. Similarly, he does not feel a gain when using the car more often than expected due to bad weather. The consumer feels a loss, however, if the rental price depends on the driven miles and he used the car more often than expected. One could defend this assumption also on the ground that there is one point in time where the consumer receives his bill and compares it with his expectations, whereas the potential gains and losses regarding the consumption of the good are distributed among the whole billing period and therefore less salient.

The consumer's gain-loss function is assumed to be piece-wise linear, since the main driver of loss aversion—in particular for small stakes—is the kink in the value function and not its diminishing sensitivity.[16] If the consumer pays T, but expected to pay $\hat{T}$, then his

[14] Strictly speaking, we have to take the left-hand limit when q approaches $q^S(\theta)$ to obtain the stated derivative.

[15] This assumption is also imposed by Spiegler (2010). The implications of not assuming a general gain-loss function for both dimensions are investigated by Karle and Peitz (2010b). They show that this assumption has merely quantitative effects.

[16] The assumption of a piece-wise linear gain-loss function is in accordance with the majority of applied loss-aversion articles. See, for instance, Heidhues and Kőszegi 2008) or Herweg et al. (forthcoming).

gain-loss utility is given by

$$\mu(\hat{T}-T)=\begin{cases}\eta(\hat{T}-T) & , \quad \text{for } \hat{T}\geq T \\ -\eta\lambda(T-\hat{T}), & \text{for } T>\hat{T}\end{cases} ,$$

where $\eta>0$ is the weight put on gain-loss utility relative to intrinsic utility and $\lambda\geq 1$ is the weight put on losses relative to gains. With $\lambda>1$ the consumer is loss averse in the sense that losses loom larger than gains of equal size. For $\eta=0$ the consumer's preferences are not reference dependent.

The consumer's expected demand conditional on the type fully determines the distribution of his expected payments, and thus his reference point. Suppose the consumer signed the contract. Then his overall utility from this contract when purchasing q units, given his demand type is ϕ and his expected consumption is $\langle q(\theta)\rangle_{\theta\in\Theta}$, is given by

$$U(q|\phi,\langle q(\theta)\rangle)=u(q,\phi)-T(q)+\eta\int_{X(q)}[T(q(\theta))-T(q)]f(\theta)\,d\theta \\ -\eta\lambda\int_{X^c(q)}[T(q)-T(q(\theta))]f(\theta)\,d\theta\ , \quad (3)$$

with $X(q)\equiv\{\theta\in\Theta|T(q)<T(q(\theta))\}$ and $X^c(q)\equiv\{\theta\in\Theta|T(q)\geq T(q(\theta))\}$. Observe that for $p\geq 0$ it holds that $X(q)\supseteq X(q+z)$ and $X^c(q)\subseteq X(q+z)$ for any $z>0$. In words, if the consumer demands a higher quantity this increases the number of demand types compared to which he feels a loss and reduces the number of demand types compared to which he feels a gain.

To deal with the resulting interdependence between actual consumption and expected consumption, we use the personal equilibrium concept, which requires the strategy that generates expectations to be optimal conditional on these expectations.[17]

Definition 1 (Personal Equilibrium) *For a given per unit price p the demand function $\langle\hat{q}(\theta;p)\rangle_{\theta\in\Theta}$ is a personal equilibrium if for all $\phi\in\Theta$,*

$$\hat{q}(\phi;p)\in \arg\max_{q\geq 0}\ U(q|\phi,\langle\hat{q}(\theta;p)\rangle)\ .$$

4. THE DEMAND FUNCTION

In this section, we characterize the consumer's demand given he accepted the two-part tariff (p,L). First, we characterize some basic properties the demand function needs to satisfy to constitute a personal equilibrium. Thereafter, it is shown that the demand function that constitutes a personal equilibrium is unique if the marginal price is not too high. For high marginal prices there are multiple personal equilibria. In these equilibria, there is an interval of demand types where demand is independent of the type. We will show, that—given marginal costs are not too high—it is never optimal for the monopolist to charge such a high marginal price that multiplicity of personal equilibria occurs. This allows us to focus on the cases where a unique personal equilibrium exists when solving the firm's tariff choice problem, even though the marginal price is an endogenous variable.

[17] See Kőszegi and Rabin (2006, 2007) for a general description and a defense of this concept of consumer behavior.

4.1. Preliminary Characterization of the Demand Function

We can restrict attention to nonnegative per-unit prices, $p \geq 0$: A negative unit prices cannot be optimal since overconsumption is harmless. In order not to render flat-rate tariffs infeasible, it is assumed that in case of being indifferent between two or more quantities the consumer chooses the lowest of these quantities. Alternatively, one could assume that overconsumption is not completely harmless.

Since a higher demand type is associated with a stronger preference for the good, it seems reasonable that the demand function is increasing in the type. The following lemma shows that this indeed is the case.

Lemma 1 *For any two demand types $\theta_1, \theta_2 \in \Theta$ with $\theta_1 < \theta_2$, $\hat{q}(\theta_1; p) \leq \hat{q}(\theta_2; p)$.*

Unless specified otherwise, all proofs are presented in the appendix. Since in any personal equilibrium demand is increasing in the demand type and $p \geq 0$, the consumer feels losses compared to lower types and gains compared to higher ones. Thus, the consumer's utility for a given type ϕ in a personal equilibrium can be written as

$$U(\hat{q}(\phi;p)|\phi, \langle\hat{q}(\theta;p)\rangle) = u(\hat{q}(\phi;p), \phi) - T(\hat{q}(\phi;p)) + \eta \int_{\phi}^{\bar{\theta}} [T(\hat{q}(\theta;p)) - T(\hat{q}(\phi;p))] f(\theta)\, d\theta - \eta\lambda \int_{\underline{\theta}}^{\phi} [T(\hat{q}(\phi;p)) - T(\hat{q}(\theta;p))] f(\theta)\, d\theta \,. \tag{4}$$

Taking the expected value with respect to the demand type of the above formula yields the consumer's expected utility on the equilibrium path,

$$\mathbb{E}_\theta[U(\hat{q}(\theta;p)|\theta, \langle q(\theta;p)\rangle)] = \int_{\underline{\theta}}^{\bar{\theta}} [u(\hat{q}(\theta;p);\theta) - T(\hat{q}(\theta;p))] f(\theta)\, d\theta - \eta(\lambda - 1) \int_{\underline{\theta}}^{\bar{\theta}} \int_{\theta}^{\bar{\theta}} [T(\hat{q}(\phi;p)) - T(\hat{q}(\theta;p))] f(\phi) f(\theta)\, d\phi d\theta \,. \tag{5}$$

The first integral of the above formula represents standard expected intrinsic utility. The second term is the ex ante expected net loss of the consumer, which is weighted by $\eta(\lambda - 1)$. Remember that the consumer compares a given outcome with each possible other outcome. Thus, the comparison of any two possible payments enters the consumer's expected utility exactly twice, once as a loss and once as an equally-sized gain. For $\lambda = 1$ the consumer puts equal weights on gains and losses, hence, ex ante all these comparisons cancel out. When making the purchasing decision, however, even for $\lambda = 1$ these comparisons do not cancel out since expectations are fixed. We will further explain this observation in the next section.

Even with loss-averse consumers demand is well behaved in the following sense: The consumer's demand when playing a personal equilibrium does not "jump" if the demand type changes slightly.

Lemma 2 *Any demand function that constitutes a personal equilibrium $\langle\hat{q}(\theta;p)\rangle_{\theta\in\Theta}$ is continuous in the demand type θ.*

4.2. A Candidate for a Personal Equilibrium Demand Function

In this part, we show that demand is strictly increasing in the type if the marginal price is relatively low. To establish this result, we define the function $\tilde{q}(\theta;p)$. In fact—as we will show—the function $\tilde{q}(\theta;p)$ characterizes the unique personal equilibrium if p is sufficiently small. The function $\tilde{q}(\theta;p)$ is implicitly characterized by

$$\frac{\partial u(\tilde{q}(\theta;p),\theta)}{\partial q} \equiv p\ [1+\eta+\eta(\lambda-1)F(\theta)]\ . \tag{6}$$

For a standard consumer with $\eta = 0$, the function $\tilde{q}(\theta;p)$ equates marginal utility with marginal payments. For a consumer with reference-dependent preferences but who is not loss averse, i.e. $\lambda = 1$, the marginal value of money at the point of the quantity decision is $1+\eta$. Paying \$1 more reduces the intrinsic utility by 1 and it reduces the gain-loss utility which is weighted by η: either by increasing the losses by 1 or by reducing the gains by 1. If the consumer also is loss averse, he additionally perceives a net loss compared to lower demand types which are paying lower bills. Thus, loss aversion leads to a downward distortion of demand, in particular for high-demand types.

It turns out that whether or not $\tilde{q}(\theta;p)$ is strictly increasing in the type plays an important role for the characterization of personal equilibria. Implicit differentiation of (6) with respect to θ yields

$$\frac{d\tilde{q}(\theta;p)}{d\theta} = -\frac{\partial^2 u(\tilde{q}(\theta;p),\theta)/\partial q\partial\theta - p\eta(\lambda-1)f(\theta)}{\partial^2 u(\tilde{q}(\theta;p),\theta)/\partial q^2}\ .$$

The function $\tilde{q}(\theta;p)$ is strictly increasing in θ if and only if the following condition is satisfied:

Condition 1 *For all $\theta \in \Theta$,*

$$p < \frac{\partial^2 u(\tilde{q}(\theta;p),\theta)/\partial q\partial\theta}{\eta(\lambda-1)f(\theta)}\ . \tag{C1}$$

The right-hand side of (C1) tends to infinity, and thus Condition 1 is always satisfied, if either η tends to zero or λ tends to 1. The condition is more likely to be satisfied if the distribution of the demand types is not very dense. In other words, if the environment is sufficiently unpredictable then Condition 1 holds. Furthermore, the above condition is satisfied if the per unit price is sufficiently small.

Now we are prepared to establish the result that in a personal equilibrium there cannot be a set of types that consumes the same amount if the per unit price is low.

Lemma 3 *In any personal equilibrium it holds: For every pair of types $\theta_1, \theta_2 \in \Theta$ with $\theta_1 \neq \theta_2$ it holds that $\hat{q}(\theta_1;p) \neq \hat{q}(\theta_2;p)$ if and only if Condition 1 holds.*

Unique Personal Equilibrium.—The demand function $\hat{q}(\theta;p)$ is nondecreasing and continuous and thus differentiable with respect to q for $q \in [\hat{q}(\underline{\theta};p), \hat{q}(\bar{\theta};p)]$ almost everywhere. Given Condition 1 holds, the personal equilibrium is strictly increasing and thus we can conclude that $d\hat{q}(\theta;p)/d\theta > 0$.

In the following, we derive necessary conditions a personal equilibrium demand function has to satisfy if it is strictly increasing in the demand type. It is shown that there exists exactly one candidate satisfying the necessary conditions. This unique candidate is indeed strictly increasing if and only if Condition 1 holds. We proceed by construction, that is, we assume the consumer expects that his demand is strictly increasing in the demand type. The utility of a consumer of type ϕ who consumes $q \in [\hat{q}(\underline{\theta};p), \hat{q}(\bar{\theta};p)]$ units, given he expected to play a personal equilibrium where consumption is strictly higher for higher types, is given by

$$U(q|\phi, \langle \hat{q}(\theta;p) \rangle) = u(q;\phi) - pq - L + \eta p \int_{\alpha(q)}^{\bar{\theta}} [\hat{q}(\theta;p) - q] f(\theta)\, d\theta - \eta\lambda p \int_{\underline{\theta}}^{\alpha(q)} [q - \hat{q}(\theta;p)] f(\theta)\, d\theta \ , \quad (7)$$

where $\alpha(q)$ is implicitly defined by $\hat{q}(\alpha(q);p) \equiv q$. Note that the derivative $\alpha'(q) = (d\hat{q}(\alpha(q);p)/d\theta)^{-1} > 0$ almost everywhere by hypothesis. Taking the derivative of $U(q|\cdot)$ with respect to q yields

$$\frac{dU(q|\cdot)}{dq} = \frac{\partial u(q,\phi)}{\partial q} - p - \eta\lambda p[q - \hat{q}(\alpha(q);p)] f(\alpha(q))\alpha'(q) - \eta\lambda p \int_{\underline{\theta}}^{\alpha(q)} f(\theta)\, d\theta - \eta p[\hat{q}(\alpha(q);p) - q] f(\alpha(q))\alpha'(q) - \eta p \int_{\alpha(q)}^{\bar{\theta}} f(\theta)\, d\theta \ . \quad (8)$$

Taking into account that $\hat{q}(\alpha(q);p) - q = 0$, the above derivative can be simplified to

$$\frac{dU(q|\cdot)}{dq} = \frac{\partial u(q,\phi)}{\partial q} - p - p\eta[1 - F(\alpha(q))] - p\eta\lambda F(\alpha(q)) \ .$$

The consumer's utility is strictly concave for $q \in [\hat{q}(\underline{\theta};p), \hat{q}(\bar{\theta};p)]$, since

$$\frac{d^2U(q|\cdot)}{dq^2} = \frac{\partial^2 u(q,\phi)}{\partial q^2} - p\eta(\lambda - 1) f(\alpha(q))\alpha'(q) < 0 \ .$$

A necessary condition for $\langle \hat{q}(\theta;p) \rangle_{\theta\in\Theta}$ to constitute a personal equilibrium that is strictly increasing in the demand type is that for all $\theta \in \Theta$ the first-order condition $dU(\hat{q}(\theta;p)|\theta,\cdot)/dq = 0$ is satisfied. Thus, a necessary condition for a personal equilibrium with a strictly increasing demand function is that it satisfies for all $\theta \in \Theta$ the first-order condition (6). In other words, $\hat{q}(\theta;p) \equiv \tilde{q}(\theta;p)$. Note that the first-order condition (6) characterizes a unique candidate for a personal equilibrium demand function, that is strictly increasing in the type. We know that $d\tilde{q}(\theta;p)/dq > 0$ for all $\theta \in \Theta$ if and only if Condition 1 is satisfied. Thus, provided that Condition 1 holds the unique personal equilibrium is $\hat{q}(\theta;p) \equiv \tilde{q}(\theta;p)$. Furthermore, note that $\hat{q}(\theta;p) = q^S(\theta)$ for $p = 0$.

Proposition 1 *Suppose Condition 1 holds. Then there exists a unique personal equilibrium $\langle \hat{q}(\theta;p) \rangle_{\theta\in\Theta}$. The personal equilibrium is characterized by $\partial u(\hat{q}(\theta;p),\theta)/\partial q \equiv p[1 + \eta + \eta(\lambda - 1)F(\theta)]$.*

Since Condition 1 is satisfied for a flat-rate tariff, the next result is immediately obtained from Proposition 1.

Corollary 1 *Suppose $p = 0$, then there is a unique personal equilibrium, in which the consumer demands his satiation quantity, i.e., $\hat{q}(\theta;0) = q^S(\theta)$ for all $\theta \in \Theta$.*

Before characterizing the personal equilibria for the cases where Condition 1 fails to hold, we define $S(p)$ as the expected joint surplus of a firm and a consumer when contracting at marginal price p. Formally

$$S(p) \equiv \mathbb{E}_\theta[U(\hat{q}(\theta;p)|\theta, \langle\hat{q}(\theta;p)\rangle)] + (p-c)\mathbb{E}_\theta[\hat{q}(\theta;p)] + L \ . \tag{9}$$

Thus, the generated joint surplus from a flat-rate tariff amounts to

$$S(0) \quad = \quad \int_{\underline{\theta}}^{\bar{\theta}} \left[u(q^S(\theta),\theta) - cq^S(\theta)\right] f(\theta)\, d\theta. \tag{10}$$

The surplus generated by a flat rate, $S(0)$, becomes arbitrarily negative for sufficiently large marginal cost c and approaches the first-best surplus S^{FB} for $c \to 0$, where $S^{FB} := \max_{\langle q(\theta)\rangle_\theta \in \Theta} \int_{\underline{\theta}}^{\bar{\theta}} (u(q(\theta),\theta) - cq(\theta))\, f(\theta) d\theta$. This is intuitively plausible since the firm severely suffers from overconsumption induced by a flat-rate if marginal costs are high. For marginal costs of zero, on the other hand, the flat-rate contract implements the efficient quantities.

4.3. Personal Equilibrium with Bunching

How does the personal equilibrium look like if Condition 1 fails to hold? In this case there exists an interval of demand types for which demand is the same. Before characterizing the personal equilibrium candidates, we show that any personal equilibrium demand is bounded from below and from above. Let the lower and the upper bound be denoted by q^{MIN} and q^{MAX}, respectively. Clearly, these bounds depend on the marginal price p. It is straightforward to show that the bounds are characterized by the following equations,[18]

$$\frac{\partial u(q^{MIN},\underline{\theta})}{\partial q} = (1+\eta\lambda)p \qquad \text{and} \qquad \frac{\partial u(q^{MAX},\bar{\theta})}{\partial q} = (1+\eta)p.$$

For $q < q^{MIN}$ even the lowest type, $\underline{\theta}$, has an incentive to deviate to a higher quantity. Similarly, for $q > q^{MAX}$ it is optimal for all types, even for the highest type, $\bar{\theta}$, to deviate to a lower quantity.

By Lemma 2, any personal equilibrium is continuous in the demand type even if Condition 1 does not hold. Furthermore, if the personal equilibrium consists of flat parts as well as strictly increasing parts, then for the strictly increasing parts the personal equilibrium is given by $\tilde{q}(\theta;p)$. Thus, if the flat part is an interior interval of Θ, then at the boundary points condition (6) has to hold. On the other hand, if the flat segment starts at $\underline{\theta}$ or ends at $\bar{\theta}$, then the "bunching quantity" $\bar{q}$ has to satisfy an inequality constraint: given $\bar{q}$ a downward (upward) deviation has to reduce the utility of the type $\underline{\theta}$ (respectively $\bar{\theta}$). The following lemma characterizes these cases.

[18] Suppose $\hat{q}(\theta;p) < q^{MIN}$ for some $\theta \in \Theta$. If the consumer with demand type θ chooses a quantity $q \geq q^{MIN}$ his utility is at least $u(q,\theta) - pq - L - \eta\lambda p \int_{\underline{\theta}}^{\bar{\theta}} [q - \hat{q}(\phi;p)] f(\phi) d\phi$. (The worst case is to perceive a loss compared to all other demand types). Thus, $\partial u(q^{MIN},\theta)/\partial q - (\eta\lambda + 1)p > 0$ is a sufficient condition that type θ has an incentive to deviate to a quantity $q \geq q^{MIN}$. The lowest incentive for an upward deviation has type $\underline{\theta}$, which characterizes the bound q^{MIN}. Put verbally, quantities lower than q^{MIN} are not optimal even for the lowest demand type and even if higher quantities are accompanied with perceiving a loss compared to all other demand types. The upper bound, q^{MAX}, is obtained by a similar reasoning.

Lemma 4 *Consider a personal equilibrium $\langle \hat{q}(\theta;p)\rangle_{\theta\in\Theta}$ with bunching in at least one interval $I \subseteq \Theta$ with bounds θ_1 and θ_2 where $\underline{\theta} \le \theta_1 < \theta_2 \le \bar{\theta}$, i.e., $\hat{q}(\theta;p) = \bar{q}\ \forall\ \theta \in I$. Then the constant quantity $\bar{q}$ and the bounds, θ_1 and θ_2, are characterized by*

$$\partial u(\bar{q},\theta_1)/\partial q - p[1+\eta+\eta(\lambda-1)F(\theta_1)] = 0 \qquad \text{if } \theta_1 > \underline{\theta}$$
$$\partial u(\bar{q},\underline{\theta})/\partial q - p[1+\eta] \ge 0 \qquad \text{if } \theta_1 = \underline{\theta}$$

and

$$\partial u(\bar{q},\theta_2)/\partial q - p[1+\eta+\eta(\lambda-1)F(\theta_2)] = 0 \qquad \text{if } \theta_2 < \bar{\theta}$$
$$\partial u(\bar{q},\bar{\theta})/\partial q - p[1+\eta\lambda] \le 0 \qquad \text{if } \theta_2 = \bar{\theta}.$$

For the parts where the personal equilibrium is strictly increasing $\hat{q}(\theta;p) \equiv \tilde{q}(\theta;p)$.

The situation described in the above lemma is depicted in Figure 1.

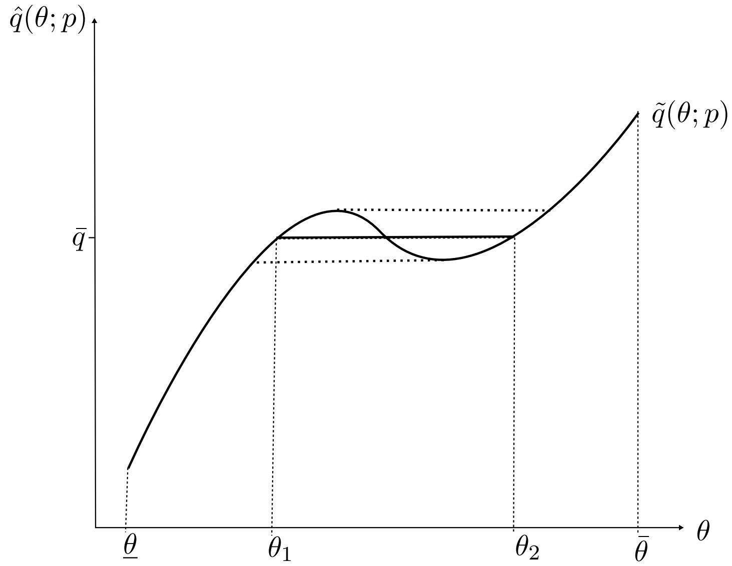

Figure 1: Personal Equilibrium with Bunching

What is the reason for demand not being strictly increasing in the type? When the personal equilibrium demand function consists of flat parts, then the personal equilibrium itself insures the consumer against fluctuations in his billing amount—at least to some degree. The consumer values this insurance if fluctuations lead to high expected losses, which is the case if p is high and if the marginal utility does not differ by too much between different types, i.e. if $\partial^2 u(q;\theta)/\partial q\partial\theta$ is low. A consumer ex ante may prefer a demand function that consists of flat parts, but to expect this function is not necessarily credible. A consumer's quantity decision depends on the realized demand type and on his expectations. A dense distribution of demand types amplifies the feedback on which the possibility that a self-fulfilling expected demand functions consist of flat parts hinges. This is also the reason why there are multiple personal equilibria, if the demand function is not strictly increasing. If, on the other hand, the demand function is not very dense then the expected demand for a given demand type has only a minor impact on the consumer's purchasing decision. In other words, for a widespread distribution of demand types only demand functions that are strictly increasing in the demand type are credible, since a higher demand type is associated with a higher (intrinsic) marginal utility.

Without further assumptions on the utility function and the distribution of the demand types, the bunching regions are intricate to characterize. Moreover, if Condition 1 does not hold and thus the personal equilibrium consists of flat parts, then there typically are multiple personal equilibria.[19] A flat rate can only be optimal if marginal costs are not too high. As we will show, if marginal costs are relatively low, then the gains from trade for high marginal prices such that bunching occurs are lower than the gains from trade generated by a flat rate. Hence, for low marginal costs it is never optimal for the monopolist to set such high per-unit prices that bunching occurs. When solving for the optimal tariff, we provide a sufficient condition that allows us to focus on the cases in which Condition 1 is satisfied. Put differently, since we are interested in situations where it is optimal for firms to offer flat-rate tariffs, there is no need to further discuss the personal equilibria for high per-unit prices.

For illustrative purposes, we characterize all personal equilibria for a special case. Suppose that types are uniformly distributed and that the cross derivative of the utility function is constant. Then, $d\tilde{q}(\theta;p)/d\theta$ is either strictly increasing for all $\theta \in \Theta$ or nonincreasing for all θ. Thus, depending on the per-unit price, the personal equilibrium is either strictly increasing or constant over all states of the world.

Corollary 2 *Suppose $\partial^2 u(q,\theta)/\partial q \partial\theta = K > 0$ for $q \leq q^S(\theta)$ and $\theta \sim U[\underline{\theta},\bar{\theta}]$. Then (i) for $p < K(\bar{\theta}-\underline{\theta})/[\eta(\lambda-1)]$ there exists a unique personal equilibrium which is characterized by $\partial u(\hat{q}(\theta;p),\theta)/\partial q \equiv p[1+\eta+\eta(\lambda-1)(\theta-\underline{\theta})/(\bar{\theta}-\underline{\theta})]$, (ii) for $p \geq K(\bar{\theta}-\underline{\theta})/[\eta(\lambda-1)]$ in any personal equilibrium demand is independent of the demand type, i.e., $\hat{q}(\theta;p)=\bar{q}$ for all $\theta \in [\underline{\theta},\bar{\theta}]$. In this case there are multiple personal equilibria and $\bar{q}$ satisfies $[\partial u(\bar{q},\bar{\theta})/\partial q]/(1+\eta\lambda) \leq p \leq [\partial u(\bar{q},\underline{\theta})/\partial q]/(1+\eta)$.*

5. The Optimality of Flat-Rate Tariffs

The monopolist maximizes its revenues minus costs subject to the consumers' participation constraint:

$$\max_{L,p\geq 0} \quad L + (p-c)\int_{\underline{\theta}}^{\bar{\theta}} \hat{q}(\theta,p) f(\theta) d\theta$$
$$\text{subject to} \quad \mathbb{E}_\theta\left[U(\hat{q}(\theta,p)|\theta,\langle \hat{q}(\phi,p)\rangle)\right] \geq 0.$$

For any unit price p, the optimal fixed fee is determined by the binding participation constraint. Thus, the monopolist's tariff choice problem can be restated as a problem of choosing only the unit price p. Since there is no asymmetric information at the contracting stage, the optimal unit price, $\hat{p}$, maximizes the joint surplus $S(p)$ of the two contracting parties.

[19] Kőszegi and Rabin (2006) define the preferred personal equilibrium as refinement for situations where multiple personal equilibria exist.

The joint surplus is given by

$$S(p) = \int_{\underline{\theta}}^{\bar{\theta}} \left\{ u(\hat{q}(\phi, p), \phi) - c\hat{q}(\phi, p) + \eta p \int_{\phi}^{\bar{\theta}} [\hat{q}(\theta, p) - \hat{q}(\phi, p)] f(\theta) \, d\theta \right. \\ \left. - \eta\lambda p \int_{\underline{\theta}}^{\phi} [\hat{q}(\phi, p) - \hat{q}(\theta, p)] f(\theta) \, d\theta \right\} f(\phi) \, d\phi. \quad (11)$$

Remember that a given difference in tariff payments enters the consumer's expected utility exactly twice, once as a gain and once as a loss. Adding expected gains and losses allows us to rewrite the joint surplus as follows

$$S(p) = \int_{\underline{\theta}}^{\bar{\theta}} [u(\hat{q}(\phi, p), \phi) - c\hat{q}(\phi, p)] f(\phi) \, d\phi \\ - \eta(\lambda - 1)p \int_{\underline{\theta}}^{\bar{\theta}} \int_{\phi}^{\bar{\theta}} [\hat{q}(\theta, p) - \hat{q}(\phi, p)] f(\theta) f(\phi) \, d\theta d\phi.$$

For $\lambda = 1$—the consumer is not loss averse—the joint surplus is the expected utility of the consumer minus the firm's expected costs of production. A loss averse consumer expects ex ante to bear a net loss if tariff payments depend on his demand type. This expected net loss reduces the joint surplus and in a sense captures the consumer's flat-rate bias.

Given Condition 1 does not hold then $S(p)$ is not unambiguously defined, since the personal equilibrium is not unique. Due to the next lemma, one can focus on the cases in which the personal equilibrium is unique given that the marginal cost is low. Let $\bar{p} := \min_{\theta}\{\kappa[\eta(\lambda - 1)f(\theta)]^{-1}\}$. Note that if $p < \bar{p}$ then Condition 1 is satisfied.

Lemma 5 *Suppose marginal cost, $c > 0$, is sufficiently low. Then the joint surplus, $S(p)$, is maximized for a unit price $p \in [0, \bar{p})$.*

The condition under which Lemma 5 is applicable is not very restrictive if one is prepared to assume that the distribution of the demand types is not very dense. With $f(\theta)$ being small and thus $\bar{p}$ being high, the possible gains from trade with unit prices larger than $\bar{p}$ are small, since demand is decreasing in p. Keep in mind that the price $\bar{p}$ is independent of the marginal cost. For the sake of argument suppose the consumer is not loss averse. If marginal cost is relatively low compared to the price $\bar{p}$, then unit prices $p \geq \bar{p}$ lead to higher distortions in demand than a unit price of zero compared to the efficient quantities. Thus, the monopolist would optimally choose a unit price $p \in [0, \bar{p})$. If the consumer is loss averse, a unit price $p \geq \bar{p}$ not only leads to greater distortions in demand than a flat rate, it additionally imposes an expected net loss on the consumer. When signing a flat-rate contract, in contrast, the consumer does not expect to incur a net loss. Since the firm tries to maximize the joint surplus—including gain-loss utility—the optimal unit price is below $\bar{p}$ for marginal cost not too high.

In all what follows it is assumed that c is such that Lemma 5 is applicable. Hence, we can focus on the case where the personal equilibrium $\langle \hat{q}(\theta, p) \rangle_{\theta \in \Theta}$ is characterized by $\partial u(\hat{q}(\theta, p), \theta)/\partial q \equiv p[1 + \eta + \eta(\lambda - 1)F(\theta)]$. The derivative of the joint surplus with respect to the marginal price p is

$$S'(p) = \int_{\underline{\theta}}^{\bar{\theta}} \left\{ (p-c)\frac{d\hat{q}(\theta,p)}{dp} + p\eta \int_{\theta}^{\bar{\theta}} \frac{d\hat{q}(\phi,p)}{dp} f(\phi)\, d\phi + p\eta\lambda \int_{\underline{\theta}}^{\theta} \frac{d\hat{q}(\phi,p)}{dp} f(\phi)\, d\phi \right\} f(\theta)\, d\theta - \eta(\lambda-1)\int_{\underline{\theta}}^{\bar{\theta}} \int_{\theta}^{\bar{\theta}} [\hat{q}(\phi,p) - \hat{q}(\theta,p)] f(\phi) f(\theta)\, d\phi d\theta\ . \quad (12)$$

Obviously, without reference-dependent preferences, i.e., $\eta = 0$, the joint surplus is maximized for a cost-based tariff where $p = c$. For all $\eta \geq 0$ the demand function is downward sloping,

$$\frac{d\hat{q}(\theta,p)}{dp} = \frac{1+\eta+\eta(\lambda-1)F(\theta)}{\partial^2 u(\hat{q}(\theta,p),\theta)/\partial q^2} < 0\ . \quad (13)$$

This immediately implies that for unit prices $p \geq c$ the joint surplus is strictly decreasing in p. Thus, the optimal marginal price $\hat{p} \in [0, c)$. In order to guarantee that $S(p)$ is well behaved, we need an additional assumption. In this regard, we define

$$\Psi(p) \equiv (p-c)\int_{\underline{\theta}}^{\bar{\theta}} \frac{d\hat{q}(\theta,p)}{dp} f(\theta) d\theta - \eta(\lambda-1)\int_{\underline{\theta}}^{\bar{\theta}} \int_{\theta}^{\bar{\theta}} [\hat{q}(\phi,p) - \hat{q}(\theta,p)] f(\phi) f(\theta)\, d\phi d\theta\ .$$

Assumption (A2) *For $p \in [0, c)$, $\Psi(p)$ is non-increasing in p.*

The function $\Psi(p)$ depends on the demand function that constitutes a personal equilibrium. Assumption (A2) is satisfied, for instance, if $\partial^2\hat{q}(p,\theta)/\partial p^2 \geq 0$ and $\partial^2\hat{q}(p,\theta)/\partial p \partial\theta \geq 0$. In particular, we have to rule out that a higher marginal price leads to a reduction in expected losses, which may happen due to a highly compressed demand profile. A higher unit price has two effects on the consumer's expected net losses. On the one hand, a higher unit price increases the expected net loss due to increased variations in payments for a given demand function. On the other hand, the consumer reacts to the higher unit price by choosing a more compressed demand function, which in turn reduces his expected net losses. In summary, Assumption 2 ensures that the direct effect on the net losses is always stronger than the indirect effect. To cut back on our lengthy formulas we define

$$\Sigma(\lambda) \equiv \eta(\lambda-1)\,\frac{\int_{\underline{\theta}}^{\bar{\theta}}\int_{\theta}^{\bar{\theta}} [\hat{q}(\phi,0) - \hat{q}(\theta,0)] f(\phi) f(\theta) d\phi d\theta}{-\int_{\underline{\theta}}^{\bar{\theta}} [d\hat{q}(\theta,0)/dp] f(\theta) d\theta}\ .$$

Note that $\hat{q}(\theta,p)$ does also depend on λ. Obviously, $\Sigma(1) = 0$ and $\Sigma(\lambda)$ tends to zero if η tends to zero. Moreover, it can be shown that $\Sigma(\cdot)$ is strictly increasing in λ and thus $\Sigma(\lambda) > 0$ for $\lambda > 1$.

With this notation, we are prepared to state the main result of this section.

Proposition 2 *Suppose $(A2)$ holds. Then, the monopolist optimally offers a flat-rate tariff, i.e., $\hat{p} = 0$, if and only if $\Sigma(\lambda) \geq c$. Moreover, $\Sigma'(\lambda) > 0$.*

According to Proposition 2, a flat-rate tariff is optimal when the marginal cost is sufficiently low compared to $\Sigma(\cdot)$. In other words, a flat-rate contract is optimal when the consumer is sufficiently loss averse, since $\Sigma(\cdot)$ is increasing in the consumer's degree of loss aversion. Thus, a consumer's degree of loss aversion is directly linked to the strength of his flat-rate bias. On the one hand, a flat-rate tariff eliminates losses on the side of the consumer, which in turn increases his willingness to pay for the contract. The numerator of $\Sigma(\cdot)$ is proportional to the net loss the consumer expects to bear if the firm increases the unit price slightly above zero. On the other hand, a flat-rate tariff leads to an inefficiently high level of consumption which is costly to the firm. In sum, if marginal costs are low, the positive effect due to minimized losses outweighs the negative effect on production costs due to overconsumption, and thus a flat-rate tariff is optimal. Interestingly, a flat-rate contract can be optimal only if there is enough variation in the consumer's demand. The numerator of $\Sigma(\lambda)$ is a measure for the degree of demand variation. If the consumer's demand is independent of his type—no variation in demand—then $\Sigma(\cdot)$ equals zero. The numerator of $\Sigma(\cdot)$ averages over the cumulated differences in the satiation demand for a given type to all higher types.[20] Moreover, a measured tariff is optimal if the consumer's demand reacts sensitive to price changes. The denominator measures how strong on average the consumer's demand reacts due to an increase of the unit price slightly above zero. Since a flat-rate contract leads to overconsumption which is costly, the firm has an incentive to choose a positive unit price if price increases cause sharp reductions in demand. Thus, a flat-rate tariff is more likely to be optimal when either fluctuations in demand are high, or when demand reacts relatively inelastic to price changes.

For the sake of clarity, we stated Proposition 2 under Assumption (A2). The optimality of flat-rate contracts, however, does not rely on (A2). Without imposing (A2) a flat-rate tariff is optimal when the requirement $\Sigma(\lambda) \geq c$ is replaced by the following slightly more restrictive condition: For all $p \in [0, c)$:

$$\eta(\lambda - 1)\,\frac{\int_{\underline{\theta}}^{\bar{\theta}}\int_{\theta}^{\bar{\theta}}[\hat{q}(\phi,p) - \hat{q}(\theta,p)]f(\phi)f(\theta)d\phi d\theta}{-\int_{\underline{\theta}}^{\bar{\theta}}[d\hat{q}(\theta,p)/dp]f(\theta)d\theta} \geq c\,.$$

The intuition behind the above finding is the same as the one behind Proposition 2. Since the firm's profit—or, more precisely, the joint surplus—is not necessarily quasi concave for $p \in [0, c)$, it is not sufficient for the optimality of flat rates that $S(\cdot)$ is decreasing at $p = 0$. The above condition ensures—without directly imposing quasi concavity—that the joint surplus is decreasing in the marginal price.

Example.—To illustrate the optimality of flat-rate tariffs we now discuss an example. Suppose the consumer's intrinsic utility for the good is given by $u(q,\theta) = \theta q - (1/2)q^2$ for $q \leq \theta = q^S(\theta)$ and $u(q,\theta) = (1/2)\theta^2$ otherwise. The demand types are uniformly distributed on $[2, 3]$. Let the weight put on gain-loss utility be $\eta = 1$ and suppose that the marginal costs of production are $c = 0.05$.[21] The quantities demanded in a personal

[20] Empirical studies about the flat-rate bias who support the so-called "ratio rule" often argue that a higher variance in the consumer's demand does not necessarily increase the consumer's preferences for a flat-rate option, see for instance Nunes (2000). Similarly, for a loss averse consumer an invoice profile is more risky if it has a higher average self distance (numerator of $\Sigma(\cdot)$), which does not imply a higher variance.

[21] Normalizing $\eta = 1$ is not crucial for the insights that are to be obtained. The effects of changing η are qualitatively similar to the effects of changing λ. For $\lambda = 1$ the consumer has reference-dependent preferences but is not loss averse. His utility for a given state of the world θ is then $u(q,\theta) - 2T(q) +$ constant. The consumer values money twice at the moment of the purchasing decision, since paying one dollar more reduces intrinsic utility and reduces gain-loss utility either by reducing gains or by increasing losses. At

equilibrium are characterized by Corollary 2. For $p < (\lambda - 1)^{-1}$ the demand function is strictly increasing and given by $\hat{q}(\theta, p) = \theta[1 - p(\lambda - 1)] - 2p(2 - \lambda)$. For $p \geq (\lambda - 1)^{-1}$ the consumer's demand is independent of his demand type. In this case, there are multiple personal equilibria. Here, it can easily be verified that the joint surplus is always maximized for prices below $(\lambda - 1)^{-1}$. Nevertheless, we briefly characterize the joint surplus for all p values. For $p \geq (\lambda - 1)^{-1}$ it can be shown that the preferred personal equilibrium is to demand the highest possible quantity, i.e., $\bar{q} = \max\{2(1-p), 0\}$.[22] The joint surplus, $S(p)$, is depicted below for the case $\lambda = 3$.[23] Observe that $S(p)$ is continuous at $p = (\lambda - 1)^{-1}$ which is a general feature of the model and not due to the specific example.

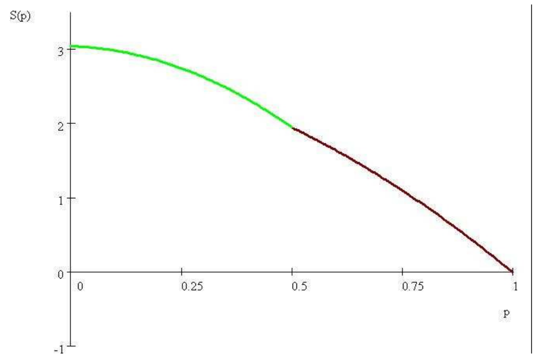

Figure 2: Joint Surplus for $\lambda = 3$.

In this example the function $\Sigma(\cdot)$ takes the following simple form, $\Sigma(\lambda) = (1/3)(\lambda - 1)(\lambda + 3)^{-1}$. Thus, by applying Proposition 2, a flat-rate tariff is optimal if $\lambda \geq 1.706$. Figure 3 depicts the joint surplus, $S(p)$, for $\lambda = 1; 1,5; 2; 3; 5$. Lower curves correspond to higher values of λ. Without loss aversion $(\lambda = 1)$ the optimal marginal price $p = (1/2)c$.

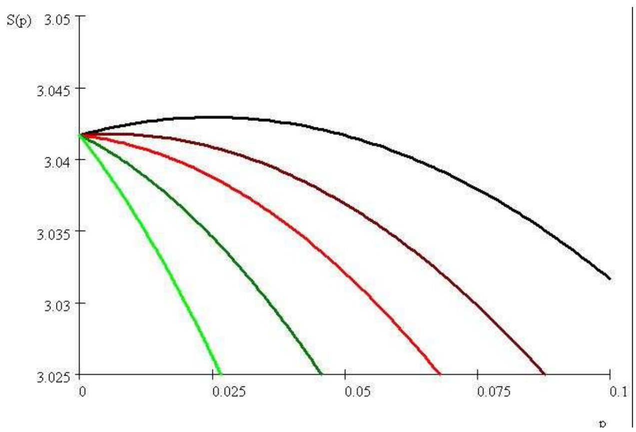

Figure 3: Joint Surplus for $\lambda = 1, 1.5, 2, 3, 5$.

Note that even for $\lambda = 1$ the consumer has reference-dependent preferences and therefore

the contracting stage, however, the consumer's expected utility for $\lambda = 1$ equals the expected utility of a consumer without gain-loss utility.

[22]The preferred personal equilibrium is the plan among the consistent plans (personal equilibria) that maximizes the consumer's expected utility. Here, the preferred personal equilibrium is also optimal from the firm's perspective.

[23]For $\eta = 1$, the conventional estimate of two-to-one loss aversion corresponds to $\lambda = 3$.

his marginal utility for money is two at the moment where he makes his purchasing decision. With the consumer being loss averse ($\lambda > 1$), in most cases a flat-rate tariff is optimal.

6. Duopolistic Competition and Heterogeneous Consumers

6.1. Market Framework

In this section, we extend the baseline model to allow for imperfect competition and heterogeneous consumers. Consider a market for one good or service where two firms, A and B, are active. Moreover, there is a continuum of ex ante heterogeneous consumers whose measure is normalized to one.

Players & Timing.—The consumers can be partitioned into two groups that differ in their degree of loss aversion. The weight put on gain-loss utility, $\eta > 0$, is the same for both groups.[24] Let the two groups be denoted by $j = 1, 2$ with $\lambda_1 < \lambda_2$. The distribution of demand types is identical for both groups of loss-averse consumers. As before, the demand type is unknown to consumers and firms at the point of contracting.

The two symmetric firms, A and B, produce at constant marginal cost $c > 0$ and without fixed cost. Each firm $i = A, B$ offers a two-part tariff to each group of consumers $j = 1, 2$. The tariff is given by $T_j^i(q) = L_j^i + p_j^i q$, where $q \geq 0$ is the quantity, and L_j^i and p_j^i denote the fixed fee and the per-unit price, respectively, charged by firm i from consumers of type j. We will analyze the symmetric information case in which firms observe λ, as well as the asymmetric information case in which λ is private information of the consumer.

The timing is as follows: (1) Firms simultaneously and independently offer a tupel of two-part tariffs $\{(L_j^i, p_j^i)\}_{j=1,2}$ to consumers. (2) Each consumer either signs exactly one contract or none. (3) Each consumer privately observes his demand type. Thereafter, each consumer who accepted a contract chooses a quantity. (4) Finally, payments are made according to the demanded quantities and the concluded contracts.

Discrete Choice Framework.—The products of the two firms are symmetrically differentiated. We assume that, next to λ, consumers are ex ante heterogeneous with respect to their brand preferences. Each consumer has idiosyncratic preferences for differing brands of the product (firms), which are parameterized by $\zeta = (\zeta^0, \zeta^A, \zeta^B)$. A consumer with brand preferences ζ has net utility $v^i + \zeta^i$ if he buys from firm i, and net utility ζ^0 if no contract is signed, where $v^i = \mathbb{E}_\theta[U(\cdot)]$. The brand preferences $\zeta = (\zeta^0, \zeta^A, \zeta^B)$ are independently and identically distributed according to a known distribution among the two groups of consumers.

To solve for the tariffs that are offered in the pure-strategy Nash equilibrium by the two firms, we follow the approach of Armstrong and Vickers (2001) and model firms as offering utility directly to consumers. Each two-part tariff can be considered as a deal of a certain expected value that is offered by a firm to its consumers. Thus, firms compete over customers by trying to offer them better deals, i.e., a two-part tariff that yields higher utility (including gain-loss utility). Put differently, we decompose a firm's problem into two parts. First, we solve for the two-part tariff that maximizes profits subject to the constraint that the consumer receives a certain utility level. Thereafter, we solve for the utility levels (v_1^i, v_2^i) a firm i of-

[24]The results would be qualitatively the same if the two groups would differ in η but not in λ.

fers to its customers. It is important to note that, when λ is unobservable, the two-part tariffs have to be designed such that each group of consumers prefers the offer that is dedicated to them. Suppose the utility offered to consumers of group j by firm A and firm B is v_j^A and v_j^B, respectively. Furthermore, assume that the incentive constraints are satisfied. Then, the market share of firm A in the submarket j is $m_j(v_j^A, v_j^B)$ and the market share of firm B is $m_j(v_j^B, v_j^A)$, with $m_j(v_j^A, v_j^B) + m_j(v_j^B, v_j^A) \leq 1$. The market share function $m_j(\cdot)$ is increasing in the first argument and decreasing in the second. Since the brand preferences are identically distributed among the two groups, the market share functions are identical for the two submarkets, i.e., $m_1(\cdot) = m_2(\cdot) = m(\cdot)$. Following Armstrong and Vickers, we impose some regularity conditions in order to guarantee existence of equilibrium. First, we assume that

$$\frac{\partial m(v^A, v^B)/\partial v^A}{m(v^A, v^B)} \text{ is non-decreasing in } v^B.$$

Second, we assume that for each submarket the collusive utility level $\tilde{v}_j$ exists which maximizes (symmetric) joint profits.[25]

6.2. Firm's Subproblem: Joint Surplus Maximization

For this part, suppose firms can observe consumers' types $\lambda \in \{\lambda_1, \lambda_2\}$. With consumers' loss aversion types being observable, the two market segments of types λ_1 and λ_2 can be viewed as distinct markets. Thus, for the analysis we can focus on one market where consumers are homogeneous with respect to their degree of loss aversion, which is denoted by λ.

Suppose firm $i \in \{A, B\}$ offers consumers a "deal" using a two-part tariff (L^i, p^i) that gives them utility v^i. Then, if a consumer with brand preferences $\zeta = (\zeta^0, \zeta^A, \zeta^B)$ purchases from firm i his net utility is $v^i + \zeta^i$. Let $\pi_j(v^i)$ be firm i's maximum profit per customer of type j when offering them a deal that yields utility v^i. The per-consumer profit function is the same for both firms but—in general—it depends on the consumer's degree of loss aversion λ. For now we focus on one market segment and therefore the subscript indicating the loss-aversion type can be omitted without confusion. Since $\pi(\cdot)$ is the same for both firms, we will omit firm's superscript in the following. With this notation, $\pi(v)$ is given by the solution to the problem:

$$\pi(v) \;=\; \max_{L,p\geq 0} \;:\; \left\{ L + (p-c)\int_{\underline{\theta}}^{\bar{\theta}} \hat{q}(\theta,p) f(\theta) d\theta \;\middle|\; \mathbb{E}_\theta[U(\hat{q}(\theta,p)|\theta, \langle q(\phi,p)\rangle)] = v \right\}. \tag{14}$$

First, we study the firm's subproblem, that is, we derive the optimal two-part tariff that solves the above problem. Thereafter, we solve for the utility levels and the corresponding tariffs which are offered by the two firms in equilibrium. Put differently, the task is to maximize a firm's profit over the choice variables p and L subject to the constraint that the consumer's expected utility from the offered deal is v. The firm's tariff choice problem can be restated as a problem of choosing only the unit price p. The firm chooses p to maximize

[25] For a detailed description of the competition-in-utility-space framework and the needed assumptions see Armstrong and Vickers (2001).

$S(p) - v$, i.e., the firm chooses the marginal price p such that the joint surplus of the two contracting parties, the consumer and the firm, is maximized. The optimal marginal price $\hat{p}$ is independent of the utility, v, the firm offers to the consumer. This immediately implies that $\pi'(v) = -1$. More importantly, the optimal marginal price is characterized by the same conditions as in the case of a monopolistic firm.

In the following we focus on the profit maximization problem of firm A. We assume that (A2) holds for both types of loss-averse consumers, i.e, for $\lambda \in \{\lambda_1, \lambda_2\}$. Moreover, it is assumed that $\Sigma(\lambda_2) \geq c$.

6.3. Symmetric Information Case

Consider market segment $j \in \{1, 2\}$. For a given utility v_j^B offered by firm B the profit maximization problem of firm A is given by

$$\max_{v_j^A} m(v_j^A, v_j^B)\pi_j(v_j^A) . \tag{15}$$

The necessary first-order condition for profit maximization amounts to

$$[\partial m(v_j^A, v_j^B)/\partial v_j^A]\pi_j(v_j^A) + m(v_j^A, v_j^B)\pi_j'(v_j^A) = 0 . \tag{16}$$

Remember that $\pi_j'(v^A) = -1$. Put verbally, the optimal marginal price is unaffected by the choice of v_j^A. If firm A offers one unit utility more to consumers, then this is optimally achieved by lowering the fixed fee by one unit. The fixed fee is a one-to-one transfer from the consumer to the firm. Define

$$\Phi(v) \equiv \frac{m(v, v)}{\partial m(v, v)/\partial v^A} .$$

Applying Proposition 1 of Armstrong and Vickers (2001), the firm's per customer profit in submarket j in the symmetric equilibrium is given by

$$\pi_j(\hat{v}_j) = \Phi(\hat{v}_j) ,$$

where $\hat{v}_j$ denotes the utility offered to consumers of type λ_j by both firms in equilibrium. As is shown by Armstrong and Vickers, there are no asymmetric equilibria. Moreover, the equilibrium often is unique.[26] The following proposition summarizes the tariffs offered by the two firms in equilibrium.

Proposition 3 (Full Information) *Suppose* $(A2)$ *holds. Then, in equilibrium, (i) if* $\Sigma(\lambda_1) < c \leq \Sigma(\lambda_2)$ *both firms offer the tariff* $(\hat{p}, \hat{L})$ *with a positive unit price to consumers of type* λ_1*, and a flat-rate tariff* $(0, L^F)$ *to consumers of type* λ_2*. (ii) If* $c \leq \Sigma(\lambda_1) < \Sigma(\lambda_2)$*, then both firms offer the flat-rate tariff* $(0, L^F)$ *to both types of loss-averse consumers. The tariffs* $(\hat{p}, \hat{L})$ *and* $(0, L^F)$ *are characterized by:* $S_1'(\hat{p}) = 0$,

$$\hat{L} = \Phi(\hat{v}_1) - (\hat{p} - c) \int_{\underline{\theta}}^{\bar{\theta}} \hat{q}_1(\theta, \hat{p}) f(\theta) d\theta$$

$$\textit{and} \quad L^F = \Phi(\hat{v}_2) + c \int_{\underline{\theta}}^{\bar{\theta}} q^S(\theta) f(\theta) d\theta ,$$

respectively, with $\hat{p} \in (0, c)$.

[26] See Armstrong and Vickers (2001) for sufficient conditions for a unique equilibrium.

If the degree of loss aversion of the less loss-averse consumers is below the threshold, then firms offer to these consumers a usage-based pricing scheme. Next to the usage-based scheme firms offer a flat-rate tariff to the more loss-averse consumers. Thus, consumer heterogeneity with respect to their first-order risk preferences provides one possible answer to the question why firms offer flat-rates next to usage-based tariffs. If the degree of loss aversion of both types is above the threshold, then firms offer only a single tariff, which is a flat-rate tariff. It is worthwhile to point out that the structure of the tariffs offered in equilibrium does not depend on the degree of competition. The degree of competition only influences the size of the basic charge. In a more competitive market firms offer tariffs with lower basic charges, whereas the unit price is unaffected by the degree of competition. Even in the limit, when we approach a perfectly competitive market, the equilibrium tariffs do not converge to marginal cost pricing. Note that, in this model, the degree of competition (or the degree of product differentiation) is measured by $\Phi(\cdot)$. A lower $\Phi(\cdot)$ corresponds to a more competitive market. $\Phi(\cdot)$ is the inverse semi-elasticity of demand evaluated at the equilibrium utility level. Thus, the higher $\Phi(\cdot)$ the less elastic is the demand of a firm. To make this point even clearer, suppose firms are located at the two extreme points of a Hotelling line of length one. Consumers' ideal brands are uniformly distributed on this line. If a consumer incurs "transport cost" of t times the distance between his ideal brand and the firm he purchases from, then $\Phi(\hat{v}) = t$—given the market is fully covered in equilibrium.

A final comment to the offered tariffs is in order: Here, firms offer a flat-rate tariff to those consumers who are willing to pay an extra amount to be insured against unexpected high bills. The flat-rate tariff, however, is not offered to exploit consumers' behavioral bias. Here, firms offer flat-rate tariffs to consumers in the cases where these tariffs also maximize the joint surplus. This is in contrast to several models with biased consumers where firms design tariffs to exploit consumers' biases, see for instance Grubb (2009) or Eliaz and Spiegler (2008).

For completeness, the following result states the equilibrium outcome for the case where consumers are not loss averse.

Corollary 3 *Suppose consumers do not exhibit loss aversion, i.e.,* $\lambda_1 = \lambda_2 = 1$. *Then, in equilibrium both firms offer the two-part tariff with marginal price* $\hat{p} = [1/(1+\eta)]c$ *and fixed fee* $\hat{L} = \Phi(\hat{v}) + [\eta/(1+\eta)]c \int_{\underline{\theta}}^{\overline{\theta}} q^{FB}(\theta)d\theta$, *where* $q^{FB}(\theta) \equiv arg \max_q\{u(q,\theta) - cq\}$. *In this case the joint surplus equals the first-best surplus,* $\hat{v} + \pi(\hat{v}) = S^{FB}$.

Without loss aversion, due to ex ante contracting, firms choose a tariff that implements the first-best allocation. Depending on the degree of competition, the first-best surplus is shared between firms and consumers. Since with $\lambda = 1$ consumers still have reference-dependent preferences, the unit price does not equal marginal cost. Due to reference-dependent preferences the consumer's marginal utility for money is $1 + \eta$. It is important to point out that reference-dependent preferences without loss aversion have only quantitative effects on the equilibrium outcome but not qualitative effects.

6.4. Asymmetric Information Case

In this subsection, we investigate the tariffs offered by the two firms when facing a screening problem, i.e., the degree of loss aversion is private information of the consumer. In order to analyze this situation, we first show that a consumer's expected utility from accepting a certain two-part tariff is decreasing in his degree of loss aversion.

Lemma 6 *Consider a two-part tariff (p, L) and suppose that Condition 1 holds. Then,*

$$\frac{d}{d\lambda}\left[\mathbb{E}_{\theta}[U(\hat{q}(\theta;p)|\theta,\langle q(\theta;p)\rangle)]\right] \leq 0.$$

For the sake of argument, suppose firms offer the tariffs as in the full information benchmark. Due to Lemma 6, consumers who are less loss averse may have an incentive to choose the tariff that is designed for the more loss-averse consumers. Note that when choosing a flat-rate tariff the consumer does neither feel a loss nor a gain. Thus, the expected utility from a flat-rate contract is independent of the consumer's degree of loss aversion. Hence, if firms offer a flat-rate tariff to the types with a high degree of loss aversion, then a consumer of type λ_1 does not necessarily benefit from choosing the tariff that is designed for consumers of type λ_2. If firms' profits from the market segment of λ_1 types is lower than their profits from λ_2 types, however, then consumers of type λ_2 may have an incentive to choose the tariff $(\hat{p}, \hat{L})$. Because, if this is the case, then $\hat{v}_1$ is considerably larger than $\hat{v}_2$. We rule this out by assuming that $\Phi(\cdot)$ is non-decreasing.[27] With this assumption, both firms' profits and consumers' surplus increase in equilibrium, if the joint surplus from contracting increases. Assuming that an increase in joint surplus is shared between consumers and firms seems to be natural for imperfectly competitive markets.

With this assumption, the two types of loss-averse consumers may not exert any informational externality on each other. In other words, if this is the case, firms can screen the consumer's loss-aversion type at no cost.

Proposition 4 (Asymmetric Information) *Suppose $(A2)$ holds and that $\Phi'(v) \geq 0$. Then, (i) if $\Sigma(\lambda_1) < c \leq \Sigma(\lambda_2)$ both firms offering tariff $(\hat{p}, \hat{L})$ with a positive unit price to consumers of type λ_1, and flat-rate tariff $(0, L^F)$ to consumers of type λ_2 is an equilibrium. (ii) If $c \leq \Sigma(\lambda_1) < \Sigma(\lambda_2)$, then in equilibrium both firms offer the flat-rate tariff $(0, L^F)$ to both types of loss-averse consumers. The tariffs, $(\hat{p}, \hat{L})$ and $(0, L^F)$, are characterized in Proposition 3.*

As in the symmetric information case, if λ_1 is below and λ_2 is above the threshold, then firms offer a usage-based pricing scheme to the less loss-averse types and a flat-rate tariff to the more loss-averse consumers. The fixed fee of the flat-rate tariff is higher than the fixed fee of the usage-based pricing scheme. In this case, we do not make any claims about the uniqueness of this equilibrium.[28] If the degree of loss aversion of both types exceeds the threshold, then we obtain a pooling equilibrium: each firm offers only a single tariff that is accepted by both types of consumers.

If $\Sigma(\lambda_2) < c$ then there is an information externality. In this case, if firms can observe λ they would offer to each type a different usage-based tariff. When offering these tariffs in the asymmetric information case, then type λ_1 obtains a higher expected utility from signing the

[27] For instance, this assumption is satisfied for the standard Hotelling model and the logit demand model, see Appendix B.

[28] To analyze all equilibria we cannot apply the competition in utility space framework, since we have to take the sorting constraints explicitly into account. Note that each firm has 4 choice variables which makes the calculation of firm A's best response to firm B's tariff offers intricate.

contract that is designed for the types λ_2. We refrain from characterizing the equilibrium tariffs for this case, since this case is intricate to analyze in the applied competition-in-utility-space framework.

It is important to point out that offering a flat-rate tariff next to usage based tariffs does not impose some additional incentive constraints. If the degree of loss aversion, λ, is continuously distributed on $[\underline{\lambda}, \bar{\lambda}]$ with $\Sigma(\underline{\lambda}) < c < \Sigma(\bar{\lambda})$, then in equilibrium firms offer the flat-rate tariff $(0, L^F)$ which is chosen at least by types $\lambda \in [\tilde{\lambda}, \bar{\lambda}]$, with $\Sigma(\tilde{\lambda}) = c$.

7. Evidence and Related Literature

7.1. Existence and Causes of Tariff-Choice Biases

The existence of tariff-choice biases was first documented for U.S. households among telephone service options. Train et al. (1987) provide evidence for U.S. households favoring flat-rate tariffs over measured services for local telephone calls. Conducting a logit model with a tariff specific constant, the authors find that this constant is highest for the flat-rate option. Similar results are obtained by Train et al. (1989). The authors argue that consumers choose a tariff that ends up not being cost-minimizing for the customer's level of consumption because consumers also care about the insurance provided by the tariff option. Given uncertain consumption patterns "the flat-rate tariff provides complete insurance" (Train et al., 1989). A tendency of households to prefer flat-rate tariffs for telephone services is also reported by Hobson and Spady (1988) for single-person households, by Kling and van der Ploeg (1990) who evaluate a tariff experiment of AT&T, by Mitchell and Vogelsang (1991), and Kridel et al. (1993). For instance, Kridel et al. (1993) find that 55% of all customers who choose a flat-rate service would have achieved higher surplus if they had chosen a measured service instead. The authors also hold an insurance motive of the customers responsible for this finding. They report that customers exhibit substantial risk aversion when faced with bill uncertainty. Miravete (2003) rejects the thesis that customers are subject to a flat-rate bias when selecting telephone service tariffs. In his data set, however, the flat-rate option is optimal for the vast majority of households and thus a flat-rate bias is hard to detect.[29]

The flat-rate bias is documented also for other telecommunication services. Lambrecht and Skiera (2006) analyze transactional data of over 10,000 customers of an Internet service provider in Germany. They find that over 50% of these customers are biased in favor of a flat-rate option. Moreover, they provide evidence that the flat-rate bias is at least partially due to an insurance motive of the consumers. In a follow up paper, Lambrecht et al. (2007) argue that "[c]onsumers may prefer a tariff that leads to fewer month-to-month fluctuations in their bill". For mobile telephone services, a preference for flat-rate tariffs that cannot be explained by customers' usage is documented by Gerpot (2009) and Mitomo et al. (2009). These papers, however, rely on survey data.

[29] Based on the data set of an tariff experiment conducted by South Central Bell in 1984, Miravete (2003) finds that only 6–12% of the customers enrolled in a flat-rate contract would have saved money with the measured option. Whereas, 62–67% of those customers who selected the measured option would have saved money with the flat-rate contract. The majority of customers, 71%, however, selected the flat-rate option and this was in most of the cases optimal from a cost savings perspective. Thus, the data set of Miravete (2003) has little power to reject the flat-rate bias hypothesis.

Also relying on survey data, Nunes (2000) finds strong evidence for a flat-rate bias outside telecommunications services (grocery shopping online, access payment for a swimming pool of an apartment building). Presumably the most powerful demonstration of the flat-rate bias outside the telecommunications service sector is DellaVigna and Malmendier (2006). They analyze a data set from three U.S. health clubs and show that a large fraction of health club members who are enrolled in a flat-fee contract (either monthly or annually) paid on average more per visit than they would have paid with a pay-per-visit option. According to the authors, the leading explanation for these observations is consumers' overconfidence about future self-control. A low per usage price is a commitment device for higher attendance in case of self-control problems, when consumption leads to immediate costs and delayed rewards. Such motives of selecting an option that provides commitment to higher usage rates obviously cannot explain the prevalence of flat rates for telecommunications services, car rental, car leasing and amusement parks.

7.2. (Behavioral) Models of Pricing Strategies

Since Oi's (1971) analysis of an optimal two-part tariff for a monopolist, this pricing scheme is intensely analyzed in the economic literature. Leland and Meyer (1976) show that a firm, regardless of its objective, always does at least as well with a two-part tariff as with linear pricing. Pareto-optimal menus of tariffs are analyzed by Willig (1978). He shows that a Pareto-optimal menu includes a cost-based two-part tariff. The pricing literature of the 80's solves for the optimal nonlinear tariff. Notable works on this topic are Mussa and Rosen (1978), Maskin and Riley (1984), Goldman et al. (1984), as well as the book by Wilson (1993). This literature established the now well-known no-distortion-at-the-top result, i.e., marginal prices exceed marginal costs for all but the last unit. While these classic screening models focus on deterministic demand, there are some papers analyzing sequential screening problems. In these papers, a consumer first chooses a contract and then he learns his true preferences before making a quantity choice. See, for instance, Courty and Li (2000) or Miravete (2002).

This paper is more related to the recent and growing literature investigating how rational firms respond to consumer biases. A seminal contribution in this field is DellaVigna and Malmendier (2004). They consider a market, either monopolistic or perfectly competitive, with homogeneous time-inconsistent consumers. Their main finding is that the unit price of the optimal two-part tariff is above marginal cost for leisure goods (usage of rental car) and below marginal cost for investment goods (health club attendance). Likewise presuming that consumers are quasi-hyperbolic discounters, Heidhues and Kőszegi (forthcoming) set up a model of a perfectly competitive market for credit-cards. They allow for consumer heterogeneity and pay particularly attention to welfare implications of possible policy interventions. Using a different notion of time-inconsistency, Eliaz and Spiegler (2006) solve for the optimal menu of tariffs for a monopolist who faces consumers that differ in their degree of sophistication. The optimal contract exploits those consumers who are sufficiently naive about their self-control problems. Moreover, they show that the optimal menu can be implemented by a menu of three-part tariffs. The optimal menu does not include a flat-rate tariff. The optimal nonlinear pricing scheme for a monopolist who sells to consumers with self-control problems is also analyzed by Esteban et al. (2007). Instead of assuming time inconsistency, they model self-control problems by applying the concept of Gul and Pesendorfer (2001). The optimal tariff resembles the one in the standard nonlinear pricing literature except for a price ceiling. Similar results are obtained by Esteban and Miyagawa (2006) for a perfectly competitive market where consumers have temptation preferences according to Gul-Pesendorfer.

Next to time inconsistency, there are a few papers dealing with the optimal selling strategy for overconfident consumers. Grubb (2009) analyzes the optimal menu of nonlinear price schedules for a monopolist as well as for a perfectly competitive market. Consumers in his model are overconfident in the sense that they underestimate fluctuations in their demand. The optimal menu is close to a menu of three-part tariffs which is often observed in the cellular phone service industry. A similar model where firms screen consumers at the basis of their priors is considered by Uthemann (2005). In his model firms are differentiated à la Hotelling. Unlike Grubb, he does not assume that consumption is satiated at a finite level, and therefore he obtains that marginal prices are always above marginal cost. Focusing on only two-states of the world but without imposing any differentiability assumptions on the consumer's utility function, Eliaz and Spiegler (2008) analyze the problem of a monopolist who faces consumers with biased beliefs regarding the probability assignment to the two states of nature. Optimistic consumers, who assign too much weight on the state of nature that is characterized by larger gains from trade, sign exploitative contracts. In a stylized example, the authors show that the optimal menu may include a flat-rate tariff. The authors, however, do not derive conditions under which there model predicts flat-rate contracts.

To the best of our knowledge, there is no paper analyzing nonlinear tariffs when consumers are loss averse. Nevertheless, loss aversion and in particular the concept developed by Kőszegi and Rabin (2006, 2007) is used in models of industrial organization.[30] Heidhues and Kőszegi (2005) apply this concept to provide an explanation why monopoly prices react less sensitive to cost shocks than predicted by orthodox theory.[31] Moreover, Heidhues and Kőszegi (2008) introduce consumer loss aversion into a model of horizontally differentiated firms. They show that in equilibrium asymmetric competitors charge identical focal prices for differentiated products.[32] Next to industrial organization, the Kőszegi and Rabin formulation is applied to contract theory by Herweg et al. (forthcoming). Considering a moral hazard framework, they provide an explanation for the frequent usage of lump-sum bonus contracts. Alike considering an agency model, Macera (2009) provides a rational for creating incentives solely based on an annual performance measure even if for instance monthly performance measures are available.

8. Concluding Remarks

The goal of this article is to provide one possible explanation for the frequent usage of flat-rate tariffs. Since empirical evidence suggests that consumers choose flat rates because these

[30]Risk preferences—in particular of higher order—of decision makers that are expectation-based loss averse according to Kőszegi ad Rabin are investigated by Maier and Rüger (2009).

[31]A similar finding, in a slightly simpler setting, is obtained by Spiegler (2010).

[32]Consumer loss aversion is introduced in a model of product differentiation also bei Karle and Peitz (2010a, 2010b). Unlike Heidhues and Kőszegi, in their model consumers observe prices before making plans which product to purchase.

tariffs provide insurance in case of uncertain consumption patterns, we posit that consumers are first-order risk averse. First-order risk aversion is captured by reference-dependent preferences of the consumer in combination with loss aversion. This paper shows that offering a flat-rate contract is optimal when (i) consumers are loss averse, (ii) marginal production costs are small, and (iii) demand is uncertain. Moreover, firms offer flat-rate tariffs to those consumers whose degree of loss aversion exceeds a certain threshold. Consumers with a lower degree of loss aversion sign a metered tariff in equilibrium. Interestingly, offering a flat-rate contract next to usage-based pricing schemes does not introduce additional sorting constraints into a firm's optimization problem. Thus, this paper predicts that in markets with low marginal cost and uncertain consumption patterns, a firm's tariff menu includes a flat-rate option.

We departed from the Kőszegi-Rabin concept by positing that the consumer does not feel any sensations of gains and losses in the good dimension. If one considers gain-loss utility in both dimensions and assumes that higher demand types are always associated with higher utility, then a flat-rate tariff eliminates only the losses in the money but not in the good dimension. Depending on the particular form of the intrinsic utility function, a flat-rate contract may increase or decrease the expected losses in the good dimension. Alternatively, one could assume that intrinsic utility evaluated at the satiation quantity is constant for all demand types, but marginal utility is still increasing in the type. With this formulation, for a given type ϕ the consumer feels a loss in the good as well as in the money dimension compared to lower types $\theta < \phi$. In this case, a flat rate tariff eliminates any losses in both the good and the money dimension. Moreover, with this specification the consumer's expected utility from signing the flat-rate is independent of the degree of loss aversion. Hence, the screening result would also be robust. In summary, the formulation of this paper can be viewed as an intermediate case between the two possible approaches with gain-loss utility in both dimensions. Moreover, focusing on the case with gain-loss utility only in the money dimension helps to make the analysis of the personal equilibria clearer and shorter.

An obvious drawback of our model is that firms are restricted to two-part tariffs. With the consumer being loss averse according to Kőszegi-Rabin, his utility for a given demand type also depends on his payments for all other types. Thus, the standard procedure of the nonlinear pricing literature, where the tariff in the firm's objective typically is replaced by the consumer's net surplus, does not work here. This in turn makes the analysis of nonlinear tariffs more complicated. We believe, however, that focusing on two-part tariffs provides some insights on the forces at play when consumers are loss averse. In particular, the identified insurance motive of loss averse consumers should also play a major role when firms can offer more sophisticated contracts.

It is rather obvious that imposing a quantity limit on the flat-rate option can improve the joint surplus. If the quantity limit equals the first-best quantity for the highest state, then—compared to a flat rate with unlimited usage—standard efficiency is improved without imposing additional losses on the consumer. Flat-rate tariffs with limited usage are often observed for the Internet service industry. Hence, investigation of optimal nonlinear pricing schedules for firms facing loss-averse consumers is an interesting question for future research.

APPENDIX

A. Proofs of Propositions and Lemmas

Proof of Lemma 1: To reduce notation, we omit that demand depends on the marginal price p. Suppose, in contradiction, that $\phi_1 < \phi_2$ but $\hat{q}(\phi_1) > \hat{q}(\phi_2)$. By revealed preferences the two inequalities below follow immediately,

$$\begin{aligned} u(\hat{q}(\phi_1);\phi_1) - T(\hat{q}(\phi_1)) + \eta \int_{X(\hat{q}(\phi_1))} [T(\hat{q}(\theta)) - T(\hat{q}(\phi_1))] f(\theta)\, d\theta \\ - \eta\lambda \int_{X^c(\hat{q}(\phi_1))} [T(\hat{q}(\phi_1)) - T(\hat{q}(\theta))] f(\theta)\, d\theta \;\geq\; u(\hat{q}(\phi_2);\phi_1) - T(\hat{q}(\phi_2)) \\ + \eta \int_{X(\hat{q}(\phi_2))} [T(\hat{q}(\theta)) - T(\hat{q}(\phi_2))] f(\theta)\, d\theta - \eta\lambda \int_{X^c(\hat{q}(\phi_2))} [T(\hat{q}(\phi_2)) - T(\hat{q}(\theta))] f(\theta)\, d\theta, \end{aligned} \tag{A.1}$$

and

$$\begin{aligned} u(\hat{q}(\phi_2);\phi_2) - T(\hat{q}(\phi_2)) + \eta \int_{X(\hat{q}(\phi_2))} [T(\hat{q}(\theta)) - T(\hat{q}(\phi_2))] f(\theta)\, d\theta \\ - \eta\lambda \int_{X^c(\hat{q}(\phi_2))} [T(\hat{q}(\phi_2)) - T(\hat{q}(\theta))] f(\theta)\, d\theta \;\geq\; u(\hat{q}(\phi_1);\phi_2) - T(\hat{q}(\phi_1)) \\ + \eta \int_{X(\hat{q}(\phi_1))} [T(\hat{q}(\theta)) - T(\hat{q}(\phi_1))] f(\theta)\, d\theta - \eta\lambda \int_{X^c(\hat{q}(\phi_1))} [T(\hat{q}(\phi_1)) - T(\hat{q}(\theta))] f(\theta)\, d\theta\ . \end{aligned} \tag{A.2}$$

Subtracting (A.1) from (A.2) and rearranging yields

$$\begin{aligned} [u(\hat{q}(\phi_1);\phi_1) - u(\hat{q}(\phi_2);\phi_1)] - [u(\hat{q}(\phi_1);\phi_2) - u(\hat{q}(\phi_2);\phi_2)] &\geq 0 \\ \iff \int_{\hat{q}(\phi_2)}^{\hat{q}(\phi_1)} \frac{\partial u(q,\phi_1)}{\partial q}\, dq - \int_{\hat{q}(\phi_2)}^{\hat{q}(\phi_1)} \frac{\partial u(q,\phi_2)}{\partial q}\, dq &\geq 0 \\ \iff \int_{\hat{q}(\phi_2)}^{\hat{q}(\phi_1)} \int_{\phi_1}^{\phi_2} \frac{\partial^2 u(q,\theta)}{\partial q \partial \theta}\, d\theta dq &\leq 0\ . \end{aligned}$$

The last inequality cannot hold, since $\partial^2 u(q,\theta)/\partial q\partial\theta > 0$ for $q \leq q^S(\theta)$ by assumption and $\phi_1 < \phi_2$ and $\hat{q}(\phi_1) > \hat{q}(\phi_2)$ by hypothesis.

Q.E.D.

Proof of Lemma 2: Suppose, in contradiction, there is a personal equilibrium that is at least at one point $\phi \in \Theta$ discontinuous. If the personal equilibrium is discontinuous at ϕ then either $\hat{q}(\phi;p) < \lim_{\varepsilon\to 0} \hat{q}(\phi + |\varepsilon|;p)$ or $\lim_{\varepsilon\to 0} \hat{q}(\phi - |\varepsilon|;p) < \hat{q}(\phi;p)$. While we proof explicitly only the former case, the latter one proceeds by analogous steps. Let $\hat{q}(\phi;p) =: q_1$ and $\lim_{\varepsilon\to 0} \hat{q}(\phi + |\varepsilon|;p) =: q_2$ with $q_1 < q_2$ by discontinuity and monotonicity. First, consider a type $\theta \leq \phi$ who deviates from $\hat{q}(\theta;p) \leq q_1$ to a higher quantity $q \in (q_1, q_2)$. The utility of this type is then given by

$$U(q|\theta,\cdot) = u(q,\theta) - pq - L + p\eta \int_{\phi}^{\bar{\theta}} (\hat{q}(z;p) - q)\, f(z) dz - p\eta\lambda \int_{\underline{\theta}}^{\phi} (q - \hat{q}(z;p))\, f(z) dz\ . \tag{A.3}$$

For $q \in (q_1, q_2)$ the derivative of type θ's utility with respect to his demand is

$$\frac{dU(q|\theta,\cdot)}{dq} = \frac{\partial u(q,\theta)}{\partial q} - p[1 + \eta + \eta(\lambda - 1)F(\phi)]\ . \tag{A.4}$$

Thus, $U(q|\phi, \langle\hat{q}(\theta;p)\rangle)$ is continuous for all $q \in (q_1, q_2)$. Therefore, it has to hold that $dU/dq|_{q=q_1} \leq 0$ since $\langle\hat{q}(z;p)\rangle$ is a personal equilibrium which implies that type θ has no incentive to demand a quantity $q \in (q_1, q_2)$. Hence, the following inequality has to be satisfied

$$\partial u(q_1,\theta)/\partial q - p[1 + \eta + \eta(\lambda - 1)F(\phi)] \leq 0\ . \tag{A.5}$$

Since $\partial u(q_1,\theta)/\partial q$ is increasing in θ, the above inequality is satisfied for all types $\theta \in [\underline{\theta}, \phi]$ if it is satisfied for ϕ. Thus, it has to hold that

$$\partial u(q_1,\phi)/\partial q - p[1 + \eta + \eta(\lambda - 1)F(\phi)] \leq 0\ . \tag{A.6}$$

Note that (A.6) gives us a lower bound for q_1.

Now, consider a type $\theta > \phi$ who deviates from $\hat{q}(\theta;p) \geq q_2$ to a lower quantity $q \in (q_1, q_2)$. The marginal change in type θ's utility due to an increase in q amounts to

$$\frac{dU(q|\theta,\cdot)}{dq} = \frac{\partial u(q,\theta)}{\partial q} - p[1 + \eta + \eta(\lambda - 1)F(\phi)]\ . \tag{A.7}$$

This downward deviation is not profitable if $dU/dq|_{q=q_2} \geq 0$. Note that $\hat{q}(\theta;p) \geq q_2$ for $\theta > \phi$. Thus, the following inequality needs to be satisfied

$$\partial u(q_2,\theta)/\partial q - p[1+\eta+\eta(\lambda-1)F(\phi)] \geq 0 \; . \tag{A.8}$$

The above inequality is satisfied for all types $\theta \in (\phi, \bar{\theta}]$ if it is satisfied for type ϕ. Thus, it has to hold that

$$\partial u(q_2,\phi)/\partial q - p[1+\eta+\eta(\lambda-1)F(\phi)] \geq 0 \; . \tag{A.9}$$

The inequality (A.9) provides an upper bound for q_2. Combining inequalities (A.6) and (A.9) yields

$$\frac{\partial u(q_2,\phi)}{\partial q} \geq \frac{\partial u(q_1,\phi)}{\partial q} \; , \tag{A.10}$$

which implies that $q_2 \leq q_1$ a contradiction to $q_1 < q_2$. Thus, the demand profile of any personal equilibrium is continuous in the demand type.

Q.E.D.

Proof of Lemma 3: The sufficiency part is proved by showing that there cannot exist an interval $I \subseteq \Theta$ such that for all $\theta \in I$, $\hat{q}(\theta;p) = \bar{q}$ if Condition 1 holds. We show that there is at least one type $\hat{\theta} \in I$ who can profitably deviate to a slightly higher or slightly lower quantity than $\bar{q}$. First, the upward deviation is analyzed.

Consider a type $\hat{\theta} \in I$ who consumes $\bar{q}+\varepsilon$, with $\varepsilon > 0$ but close to zero. The utility of this type is given by

$$\begin{aligned} &U(\bar{q}+\varepsilon|\hat{\theta},\cdot) = u(\bar{q}+\varepsilon,\hat{\theta}) - p(\bar{q}+\varepsilon) - L \\ &+ \eta p \int_{\{\theta\in\Theta|\hat{q}(\theta;p)>\bar{q}+\varepsilon\}} (\hat{q}(\theta;p)-\bar{q}-\varepsilon) f(\theta)\, d\theta - \eta\lambda p \int_{\{\theta\in\Theta|\hat{q}(\theta;p)<\bar{q}+\varepsilon\}} (\bar{q}+\varepsilon-\hat{q}(\theta;p)) f(\theta)\, d\theta \; . \end{aligned}$$

Let $\Theta_L \equiv \{\theta \in \Theta | \theta < \inf\{I\}\}$ and $\Theta_H \equiv \{\theta \in \Theta | \theta > \sup\{I\}\}$. Thus, since demand is (weakly) increasing, it follows that for $\varepsilon \to 0$ it holds that $\{\theta \in \Theta | \hat{q}(\theta;p) > \bar{q}+\varepsilon\} = \Theta_H$ and $\{\theta \in \Theta | \hat{q}(\theta;p) < \bar{q}+\varepsilon\} = \Theta_L \cup I$. The increase in utility from consuming slightly more than $\bar{q}$ is

$$\begin{aligned} \left.\frac{dU(\bar{q}+\varepsilon|\hat{\theta},\cdot)}{d\varepsilon}\right|_{\varepsilon=0} &= \frac{\partial u(\bar{q},\hat{\theta})}{\partial q} - p - \eta p \int_{\theta\in\Theta_H} f(\theta)d\theta - \eta\lambda p \int_{\theta\in\Theta_L\cup I} f(\theta)d\theta \\ &= \partial u(\bar{q},\hat{\theta})/\partial q - p[1+\eta+\eta(\lambda-1)F(\theta_H)] \; , \end{aligned} \tag{A.11}$$

where $\theta_H := \inf\{\Theta_H\}$.

Next, the case of a downward deviation is considered. Utility of a type $\hat{\theta} \in I$, who consumes $\bar{q}-\varepsilon$ with $\varepsilon > 0$ is

$$\begin{aligned} &U(\bar{q}-\varepsilon|\hat{\theta},\cdot) = u(\bar{q}-\varepsilon,\hat{\theta}) - p(\bar{q}-\varepsilon) - L \\ &+ \eta p \int_{\{\theta\in\Theta|\hat{q}(\theta;p)>\bar{q}-\varepsilon\}} (\hat{q}(\theta;p)-\bar{q}+\varepsilon) f(\theta)\, d\theta - \eta\lambda p \int_{\{\theta\in\Theta|\hat{q}(\theta;p)<\bar{q}-\varepsilon\}} (\bar{q}-\varepsilon-\hat{q}(\theta;p)) f(\theta)\, d\theta \; . \end{aligned}$$

The change in utility from an infinitesimal downward deviation is given by

$$\begin{aligned} \left.\frac{dU(\bar{q}-\varepsilon|\hat{\theta},\cdot)}{d\varepsilon}\right|_{\varepsilon=0} &= -\frac{\partial u(\bar{q},\hat{\theta})}{\partial q} + p + \eta p \int_{\theta\in\Theta_H\cup I} f(\theta)d\theta + p\eta\lambda \int_{\theta\in\Theta_L} f(\theta)d\theta \\ &= -\partial u(\bar{q},\hat{\theta})/\partial q + p[1+\eta+\eta(\lambda-1)F(\theta_L)] \; , \end{aligned} \tag{A.12}$$

where $\theta_L := \sup\{\Theta_L\}$.

A deviation is not profitable if for all $\theta \in I$ it holds that $dU(\bar{q}+\varepsilon|\theta)/d\varepsilon|_{\varepsilon=0} \leq 0$ and $dU(\bar{q}-\varepsilon|\theta)/d\varepsilon|_{\varepsilon=0} \leq 0$. Thus, a necessary and sufficient condition for the existence of a personal equilibrium where all $\theta \in I$ consume $\bar{q}$ is that an upward deviation is not profitable for θ_H and that a downward deviation is not profitable for θ_L. Formally, using (A.11) and (A.12), the following two inequalities have to be satisfied:

$$\partial u(\bar{q},\theta_H)/\partial q \leq p[1+\eta+\eta(\lambda-1)F(\theta_H)], \tag{A.13}$$

$$\partial u(\bar{q},\theta_L)/\partial q \geq p[1+\eta+\eta(\lambda-1)F(\theta_L)] \; . \tag{A.14}$$

Define $\tilde{q}(\theta;p)$ such that $\partial u(\tilde{q}(\theta;p),\theta)/\partial q \equiv p[1+\eta+\eta(\lambda-1)F(\theta)]$. Inequalities (A.13) and (A.14) imply that $\tilde{q}(\theta_L;p) \geq \bar{q} \geq \tilde{q}(\theta_H;p)$. By Condition 1, $d\tilde{q}(\theta;p)/d\theta > 0$. With I being an interval we have $\theta_L < \theta_H$ and thus $\tilde{q}(\theta_L;p) < \tilde{q}(\theta_H;p)$ a contradiction. This completes the sufficiency part of the proof.

Necessity, i.e., a demand function constituting a personal equilibrium that is strictly increasing in the demand type exists only if Condition 1 hold. This part is proved in the main part of the text (Unique Personal Equilibrium). In the text, it is shown that there exists a unique candidate for a personal equilibrium demand function satisfying the necessary conditions for a strictly increasing demand function. This unique candidate also fulfills the sufficient conditions if and only if Condition 1 holds. This completes the proof.

Q.E.D.

Proof of Proposition 1: First, note that (6) characterizes the personal equilibrium almost everywhere (except at kink points). Since the candidate equilibrium is continuously differentiable, we can conclude that there are no kinks in the personal equilibrium if it is strictly increasing in θ.

Remember that local deviations $q \in [\hat{q}(\underline{\theta};p), \hat{q}(\bar{\theta};p)]$ are considered in the main body of the paper. Thus, it remains to show that there is no type who can profitably deviate to a very high or very low quantity, $q < \hat{q}(\underline{\theta};p)$ or $q > \hat{q}(\bar{\theta};p)$. To verify this claim we can focus on the case where $p > 0$.

Suppose the consumer chooses a quantity $q < \hat{q}(\underline{\theta};p)$, then his utility is given by

$$U(q|\theta,\cdot) = u(q,\theta) - pq - L + p\eta \int_{\underline{\theta}}^{\bar{\theta}} [\hat{q}(\hat{\theta};p) - q] f(\hat{\theta})\, d\hat{\theta}\ .$$

The optimal quantity in this case, q^L, is characterized by

$$\frac{\partial u(q^L,\theta)}{\partial q} = (1+\eta)p\ .$$

Thus, $q^L > \hat{q}(\theta;p)$ for $\theta > \underline{\theta}$ and $q^L = \hat{q}(\theta;p)$ for $\theta = \underline{\theta}$, a contradiction.

Now, consider the case where $q > \hat{q}(\bar{\theta};p)$. Given the demand type is θ, the consumer's utility is

$$U(q|\theta,\cdot) = u(q,\theta) - pq - L - p\eta\lambda \int_{\underline{\theta}}^{\bar{\theta}} [q - \hat{q}(\hat{\theta};p)] f(\hat{\theta})\, d\hat{\theta}\ .$$

The optimal quantity in this case, q^H is characterized by

$$\frac{\partial u(q^H,\theta)}{\partial q} = (\eta\lambda + 1)p\ .$$

Note that $q^H < \hat{q}(\theta;p)$ for $\theta < \bar{\theta}$ and $q^H = \hat{q}(\theta;p)$ for $\theta = \bar{\theta}$, again a contradiction. Hence, no type has an incentive to deviate.

Q.E.D.

Proof of Corollary 1: Follows directly from the observation that the consumer's utility for an arbitrary type is independent of the expected demand for all other types. Formally, $U(\hat{q}(\phi;p)|\phi, \langle q(\theta;p)\rangle) = u(q;\phi) - L$ which is maximized for $q \geq q^S(\phi)$. By the assumption that the consumer does not overconsume, it follows immediately that demand equals $q^S(\phi)$.

Q.E.D.

Proof of Lemma 4: First, note that if the personal equilibrium is strictly increasing in θ in some interval, then in this interval $\hat{q}(\theta;p) \equiv \tilde{q}(\theta;p)$. The proof of Proposition 1 reveals that there is a unique equilibrium candidate if $d\hat{q}/d\theta > 0$.

Suppose there exists an interval $I \subseteq \Theta$ such that $\hat{q}(\theta;p) = \bar{q}$ for all $\theta \in I$. Let $\theta_A := \inf\{I\}$ and $\theta_B := \sup\{I\}$, with $\theta_A < \theta_B$. Furthermore, assume that if $\theta_A > \underline{\theta}$ ($\theta_B < \bar{\theta}$) then there exists an neighborhood $(\theta_A - \xi, \theta_A)$ (respectively $(\theta_B, \theta_B + \xi)$) for $\xi > 0$ sufficiently small, where $\hat{q}(\cdot)$ is strictly increasing. For $\hat{q}(\cdot)$ being constant for all $\theta \in I$ it has to hold that for all types $\theta \in I$, neither a downward deviation nor an upward deviation does improve the consumer's utility.

Downward Deviation ($q < \bar{q}$): Suppose $\theta_A > \underline{\theta}$ and that a consumer with type $\theta \in I$ deviates to quantity q lower than $\bar{q}$. Let $\hat{\theta}(q)$ denote the demand type for which the consumer expected to choose this quantity q. Formally, $\hat{q}(\hat{\theta}(q);p) = q$ and $\lim_{\varepsilon\to 0} \hat{\theta}(\bar{q} - |\varepsilon|) = \theta_A$. The consumer feels a gain compared to types $(\hat{\theta}(q), \bar{\theta})$ and a loss compared to types $[\underline{\theta}, \hat{\theta}(q)]$. Note, for a minor downward deviation $\hat{\theta}'(q) > 0$. The consumer's utility from a (minor) downward deviation is

$$U^D = u(q,\theta) - pq - L + \eta p \int_{\hat{\theta}(q)}^{\bar{\theta}} (\hat{q}(\phi;p) - q) f(\phi) d\phi - \eta\lambda p \int_{\underline{\theta}}^{\hat{\theta}(q)} (q - \hat{q}(\phi;p)) f(\phi) d\phi\ . \qquad \text{(A.15)}$$

Differentiating the above utility function with respect to q yields

$$\frac{dU^D}{dq} = \frac{\partial u(q,\theta)}{\partial q} - p\left[1 + \eta + \eta(\lambda-1)F(\hat{\theta}(q))\right]\ . \qquad \text{(A.16)}$$

A downward deviation is not utility enhancing if the right-hand side of (A.16) is non-negative. The right-hand side of (A.16) is non-negative for all $q < \bar{q}$ if it is non-negative evaluated at $q = \bar{q}$. Due to the imposed Spence-Mirrlees condition, this inequality holds for all $\theta \in I$ if it holds for θ_A. Thus, it has to hold that

$$\frac{\partial u(\bar{q},\theta_A)}{\partial q} - p\left[1 + \eta + \eta(\lambda-1)F(\theta_A)\right] \geq 0\ . \qquad \text{(A.17)}$$

Note that $\hat{q}(\theta;p)$ is continuous and defined by $\tilde{q}(\theta;p)$ for θ slightly below θ_A. Hence, for $\theta_A > \underline{\theta}$ condition (A.17) has to hold with equality.

Now suppose $\theta_A = \underline{\theta}$. It is straightforward to show that a downward deviation is not utility improving if the following condition holds

$$\frac{\partial u(\bar{q}, \underline{\theta})}{\partial q} - p(1+\eta) \geq 0 \ . \tag{A.18}$$

With similar reasonings it can be shown that a non-minor downward deviation is not utility enhancing if the above inequality or (A.17) holds.

UPWARD DEVIATION ($q > \bar{q}$): Suppose $\theta_B < \bar{\theta}$. Let $\hat{\theta}(q)$ still denote the cutoff type, i.e., the consumer feels a gain compared to types $(\hat{\theta}(q), \bar{\theta}]$ and a loss compared to types $[\underline{\theta}, \hat{\theta}(q)]$. Now $\hat{\theta}(q) > \theta_B$ and $\lim_{\varepsilon \to 0} \hat{\theta}(\bar{q} + |\varepsilon|) = \theta_B$. The consumer's utility from a (minor) upward deviation is given by

$$U^U = u(q,\theta) - pq - L + \eta p \int_{\hat{\theta}(q)}^{\bar{\theta}} (\hat{q}(\phi;p) - q) f(\phi) d\phi - \eta\lambda p \int_{\underline{\theta}}^{\hat{\theta}(q)} (q - \hat{q}(\phi;p)) f(\phi) d\phi \ . \tag{A.19}$$

The derivative of U^U with respect to q is

$$\frac{dU^U}{dq} = \frac{\partial u(q,\theta)}{\partial q} - p\left[1 + \eta + \eta(\lambda - 1)F(\hat{\theta}(q))\right] \ . \tag{A.20}$$

An upward deviation is not utility enhancing for all $\theta \in I$ if

$$\frac{\partial u(\bar{q}, \theta_B)}{\partial q} - p\left[1 + \eta + \eta(\lambda - 1)F(\theta_B)\right] \leq 0 \ . \tag{A.21}$$

Since the personal equilibrium is continuous, for $\theta_B < \bar{\theta}$ the above inequality has to hold with equality.

Suppose $\theta_B = \bar{\theta}$. In this case the consumer has no incentive to choose a quantity $q > \bar{q}$ for all types $\theta \in I$ if

$$\frac{\partial u(\bar{q}, \bar{\theta})}{\partial q} - p\left[1 + \eta\lambda\right] \leq 0 \ . \tag{A.22}$$

Q.E.D.

Proof of Lemma 5: First, note that any personal equilibrium is bounded from above by $q^{MAX}(p)$, which is implicitly defined by $\partial u(q^{MAX}, \bar{\theta})/\partial q = (1+\eta)p$. Let $q^{FB}(\theta)$ denote the first-best quantities, i.e., $\partial u(q^{FB}(\theta), \theta)/\partial q = c$. In the following, we will show that for $p \geq \bar{p}$ the joint surplus, $S(p)$, is bounded from above and that this bound is lower than $S(0)$. To establish the above claim, we define $\check{q}(\theta) := \min\{q^{FB}(\theta), q^{MAX}(\bar{p})\}$. It is important to note that there is a positive mass of types for which $\check{q}(\theta) = q^{MAX}(\bar{p})$ if $\bar{p} > c/(1+\eta)$. The joint surplus generated with a unit price $p \geq \bar{p}$ is strictly lower than

$$\check{S} = \int_{\underline{\theta}}^{\bar{\theta}} \left[u(\check{q}(\theta), \theta) - c\check{q}(\theta)\right] f(\theta) d\theta \ , \tag{A.23}$$

since with a positive unit price the consumer expects to incur some net losses. A sufficient condition for $S(p)$ being maximized by a unit price $p \in [0, \bar{p})$ is that $S(0) \geq \check{S}$ (this condition is by no means necessary). $S(0) \geq \check{S}$ is equivalent to

$$\int_{\underline{\theta}}^{\bar{\theta}} \left\{ u(q^S(\theta), \theta) - u(\check{q}(\theta), \theta) - c[q^S(\theta) - \check{q}(\theta)] \right\} f(\theta) d\theta \geq 0 \ . \tag{A.24}$$

The above condition is satisfied for c being sufficiently small, which completes the proof.

Q.E.D.

Proof of Proposition 2: First, it is shown how to derive equation (12). Taking the derivative of (11) with respect to p yields

$$\begin{aligned} S'(p) = \int_{\underline{\theta}}^{\bar{\theta}} \Bigg\{ & \left(\frac{\partial u(\hat{q}(\theta,p), \theta)}{\partial q} - c \right) \frac{d\hat{q}(\theta,p)}{dp} \\ & + \eta \int_{\theta}^{\bar{\theta}} [\hat{q}(\phi,p) - \hat{q}(\theta,p)] f(\phi) \, d\phi + \eta p \int_{\theta}^{\bar{\theta}} \left[\frac{d\hat{q}(\phi,p)}{dp} - \frac{d\hat{q}(\theta,p)}{dp} \right] f(\phi) \, d\phi \\ & - \eta\lambda \int_{\underline{\theta}}^{\theta} [\hat{q}(\theta,p) - \hat{q}(\phi,p)] f(\phi) \, d\phi - \eta\lambda p \int_{\underline{\theta}}^{\theta} \left[\frac{d\hat{q}(\theta,p)}{dp} - \frac{d\hat{q}(\phi,p)}{dp} \right] f(\phi) \, d\phi \Bigg\} f(\theta) \, d\theta \ . \end{aligned} \tag{A.25}$$

The above equation can be rearranged to

$$S'(p) = \int_{\underline{\theta}}^{\bar{\theta}} \Bigg\{ \left(\frac{\partial u(\hat{q}(\theta,p),\theta)}{\partial q} - p[\eta + \eta(\lambda-1)F(\theta)] - c \right) \frac{d\hat{q}(\theta,p)}{dp} + \eta p \int_{\theta}^{\bar{\theta}} \frac{d\hat{q}(\phi,p)}{dp} f(\phi)\, d\phi + \eta\lambda p \int_{\underline{\theta}}^{\theta} \frac{d\hat{q}(\phi,p)}{dp} f(\phi)\, d\phi + \eta \int_{\theta}^{\bar{\theta}} [\hat{q}(\phi,p) - \hat{q}(\theta,p)] f(\phi)\, d\phi - \eta\lambda \int_{\underline{\theta}}^{\theta} [\hat{q}(\theta,p) - \hat{q}(\phi,p)] f(\phi)\, d\phi \Bigg\} f(\theta)\, d\theta\,. \quad (A.26)$$

Note that the following equality holds

$$\int_{\underline{\theta}}^{\bar{\theta}} \Bigg\{ \int_{\theta}^{\bar{\theta}} [\hat{q}(\phi,p) - \hat{q}(\theta,p)] f(\phi)\, d\phi - \lambda \int_{\underline{\theta}}^{\theta} [\hat{q}(\theta,p) - \hat{q}(\phi,p)] f(\phi)\, d\phi \Bigg\} f(\theta)\, d\theta = -(\lambda-1) \int_{\underline{\theta}}^{\bar{\theta}} \int_{\theta}^{\bar{\theta}} [\hat{q}(\phi,p) - \hat{q}(\theta,p)] f(\phi) f(\theta)\, d\phi d\theta. \quad (A.27)$$

Inserting (A.27) and (6) into (A.26) yields the equation (12) stated in the text. By using the definition of $\Psi(\cdot)$ the above derivative can be further simplified to

$$S'(p) = \Psi(p) + p\eta \int_{\underline{\theta}}^{\bar{\theta}} \Bigg\{ \int_{\theta}^{\bar{\theta}} \frac{d\hat{q}(\phi,p)}{dp} f(\phi) d\phi + \lambda \int_{\underline{\theta}}^{\theta} \frac{d\hat{q}(\phi,p)}{dp} f(\phi) d\phi \Bigg\} f(\theta) d\theta\,. \quad (A.28)$$

First observe that $S'(p) < 0$ for $p \geq c$. Since $\Psi(p)$ is non-increasing for $p \in [0,c)$, it holds that $S'(0) > S'(p)$ for $p \in (0,c)$. Hence, if $S'(0) \leq 0$ the joint surplus is maximized at $p = 0$. If, on the other hand, $S'(0) > 0$ then there exists a $\hat{p} \in (0,c)$ at which $S(p)$ is maximized. Since $S(\cdot)$ is continuously differentiable, the price $\hat{p}$ is characterized by the first-order condition $S'(\hat{p}) = 0$. Note, however, that the first-order condition may not be sufficient.

Next, we show that $S'(0) \leq 0$ is equivalent to $\Sigma(\lambda) \geq c$. By evaluating (12) at $p = 0$, it is obvious that $S'(0) \leq 0$ iff

$$-c \int_{\underline{\theta}}^{\bar{\theta}} \frac{d\hat{q}(\theta,0)}{dp} f(\theta) d\theta - \eta(\lambda-1) \int_{\underline{\theta}}^{\bar{\theta}} \int_{\theta}^{\bar{\theta}} [\hat{q}(\phi,0) - \hat{q}(\theta,0)] f(\phi) f(\theta)\, d\phi d\theta \leq 0\,. \quad (A.29)$$

Rearranging the above inequality and using the definition of $\Sigma(\lambda)$ reveals that $S'(0) \leq 0$ if and only if $\Sigma(\lambda) \geq c$. Finally, we verify the following claim.

Claim $\Sigma'(\lambda) > 0$.

Proof: To cut down on notation, we often write $\hat{q}(\theta)$ instead of $\hat{q}(\theta;p)$. Define $Z(\lambda)$ and $N(\lambda)$ as the numerator and the denominator, respectively, of the fraction of $\Sigma(\cdot)$. Thus,

$$Z(\lambda) \equiv \int_{\underline{\theta}}^{\bar{\theta}} \int_{\theta}^{\bar{\theta}} [\hat{q}(\phi,0) - \hat{q}(\theta,0)] f(\phi) f(\theta) d\phi d\theta, \quad (A.30)$$

and

$$N(\lambda) \equiv -\int_{\underline{\theta}}^{\bar{\theta}} [d\hat{q}(\theta,0)/dp] f(\theta) d\theta\,. \quad (A.31)$$

With this notation the derivative of $\Sigma(\cdot)$ with respect to λ can be written as

$$\Sigma'(\lambda) = \eta \frac{Z(\lambda)}{N(\lambda)} + \eta(\lambda-1) \frac{Z'(\lambda)N(\lambda) - N'(\lambda)Z(\lambda)}{N^2(\lambda)}. \quad (A.32)$$

In order to show that $\Sigma'(\lambda) > 0$, we analyze the itemized parts separately. First, we take the derivative of $\hat{q}(\cdot)$ with respect to λ which leads to

$$\frac{d\hat{q}(\cdot)}{d\lambda} = \frac{\eta p F(\theta)}{\partial^2 u(\hat{q}(\theta),\theta)/\partial q^2} \leq 0\,. \quad (A.33)$$

Thus,

$$\frac{d}{d\lambda}[\hat{q}(\phi) - \hat{q}(\theta)] = \frac{p\eta F(\phi)}{\partial^2 u(\hat{q}(\phi),\phi)/\partial q^2} - \frac{p\eta F(\theta)}{\partial^2 u(\hat{q}(\theta),\theta)/\partial q^2}\,, \quad (A.34)$$

which equals zero at $p = 0$. Hence, $Z'(\lambda) = 0$. Taking the derivative of (13) with respect to λ yields

$$\frac{d}{d\lambda}\left[\frac{d\hat{q}(\cdot)}{dp}\right] = \eta F(\theta) \left(\frac{\partial^2 u(\hat{q}(\theta),\theta)}{\partial q^2}\right)^{-1} - [1 + \eta + \eta(\lambda-1)F(\theta)] \left(\frac{\partial^2 u(\hat{q}(\theta),\theta)}{\partial q^2}\right)^{-2} \frac{\partial^3 u(\hat{q}(\theta),\theta)}{\partial q^3} \frac{d\hat{q}(\theta)}{d\lambda}\,. \quad (A.35)$$

Evaluating the above derivative at $p = 0$, and thus $d\hat{q}/d\lambda|_{p=0} = 0$, leads to

$$\frac{d}{d\lambda}\left[\frac{d\hat{q}(\cdot)}{dp}\right]\Bigg|_{p=0} = \eta F(\theta)\left(\frac{\partial^2 u(\hat{q}(\theta),\theta)}{\partial q^2}\right)^{-1} < 0 \ .$$

Thus,

$$N'(\lambda) = -\eta\int_{\underline{\theta}}^{\bar{\theta}} F(\theta)\left(\frac{\partial^2 u(\hat{q}(\theta),\theta)}{\partial q^2}\right)^{-1} f(\theta)d\theta \ .$$

Since $Z'(\lambda) = 0$, equation (A.32) simplifies to

$$\begin{aligned}\Sigma'(\lambda) &= \eta\left[\frac{Z(\lambda)}{N(\lambda)} - (\lambda-1)\frac{Z(\lambda)N'(\lambda)}{N^2(\lambda)}\right] \\ &= \eta\frac{Z(\lambda)}{N^2(\lambda)}\Big[N(\lambda) - (\lambda-1)N'(\lambda)\Big] \ . \end{aligned} \tag{A.36}$$

Since $Z(\lambda) > 0$ by $\hat{q}(\cdot)$ being non-decreasing in type, it remains to show that $N(\lambda) - (\lambda-1)N'(\lambda) > 0$, which is equivalent to

$$-\int_{\underline{\theta}}^{\bar{\theta}} [d\hat{q}(\theta,0)/dp]f(\theta)d\theta + \eta(\lambda-1)\int_{\underline{\theta}}^{\bar{\theta}} F(\theta)\left(\frac{\partial^2 u(\hat{q}(\theta),\theta)}{\partial q^2}\right)^{-1} f(\theta)d\theta > 0 \ . \tag{A.37}$$

Inserting the explicit formula, (13), for $d\hat{q}(\cdot)/dp$ into the above inequality yields

$$\begin{aligned}\int_{\underline{\theta}}^{\bar{\theta}} \Bigg\{ &- [1+\eta+\eta(\lambda-1)F(\theta)]\left(\frac{\partial^2 u(\hat{q}(\theta),\theta)}{\partial q^2}\right)^{-1} \\ &+ \eta(\lambda-1)F(\theta)\left(\frac{\partial^2 u(\hat{q}(\theta),\theta)}{\partial q^2}\right)^{-1}\Bigg\} f(\theta)d\theta > 0 \end{aligned} \tag{A.38}$$

$$\iff \int_{\underline{\theta}}^{\bar{\theta}} -(1+\eta)\left(\frac{\partial^2 u(\hat{q}(\theta),\theta)}{\partial q^2}\right)^{-1} f(\theta)d\theta > 0 \ . \tag{A.39}$$

The last inequality is satisfied since $u(\cdot)$ is a strictly concave function in q for $q \leq q^S(\theta)$. Q.E.D.

Q.E.D.

Proof of Proposition 3: In order to apply Proposition 1 of Armstrong and Vickers (2001), the following three properties have to be satisfied: (i) $[\partial m(v^A, v^B)/\partial v^A]\,[m(v^A, v^B)]^{-1}$ is non-decreasing in v^B, (ii) there exists $\tilde{v}_j > -\infty$ that maximizes $m(v,v)\pi_j(v)$ for $j = 1,2$, and (iii) for $j = 1,2$ there exists $\bar{v}_j$ defined by $\pi_j(\bar{v}_j) = 0$, $\pi_j(v) < 0$ if $v > \bar{v}_j$. Since we explicitly assumed (i) and (ii) these properties are satisfied. To see that (iii) is also satisfied note that $\bar{v}_j = \max_p\{S_j(p)\}$. Obviously, $\pi_j(\bar{v}_j) = 0$ and $\pi_j(v) < 0$ if $v > \bar{v}_j$. Hence, we can apply Proposition 1 of Armstrong and Vickers. According to this proposition there are no asymmetric equilibria and the equilibrium utility level $\hat{v}_j \in (\tilde{v}_j, \bar{v}_j)$. Since $m(v^A, v^B)\pi_j(v^A)$ is continuously differentiable, the equilibrium utility level satisfies the first-order condition of profit maximization. Thus, $\pi_j(\hat{v}_j) = \Phi(\hat{v}_j)$.

From Proposition 2 it follows that the optimal marginal price $\hat{p}_j$ is greater than zero if and only if $\Sigma(\lambda_j) < c$. If this is the case then $\hat{p}_j$ is characterized by $S'_j(\hat{p}_j) = 0$, as was shown in the proof of Proposition 2. The per customer profit of a firm is given by

$$\pi_j = L + (p-c)\int_{\underline{\theta}}^{\bar{\theta}} \hat{q}_j(\theta;p)f(\theta)d\theta \ . \tag{A.40}$$

Since, in equilibrium, $\pi_j = \Phi(\hat{v}_j)$ the equilibrium fixed fee is given by

$$L_j = \Phi(\hat{v}_j) - (p_j - c)\int_{\underline{\theta}}^{\bar{\theta}} \hat{q}_j(\theta;p_j)f(\theta)d\theta \ . \tag{A.41}$$

Replacing p_j by $\hat{p}$ and 0, leads to the fixed fees $\hat{L}$ and L^F, respectively.

Q.E.D.

Proof of Lemma 6: With slight abuse of notation, we omit for the proof that the demand function, $\hat{q}(\cdot)$, depends on the marginal price p. Define $V(\lambda;\theta)$ as the consumer's surplus for a given demand type on the equilibrium path. Formally,

$$\begin{aligned} V(\lambda;\theta) &= u(\hat{q}(\theta),\theta) - p\hat{q}(\theta) - L + \eta p\int_{\theta}^{\bar{\theta}}[\hat{q}(\phi)-\hat{q}(\theta)]f(\phi)d\phi \\ &\quad -\eta\lambda p\int_{\underline{\theta}}^{\theta}[\hat{q}(\theta)-\hat{q}(\phi)]f(\phi)d\phi \\ &= u(\hat{q}(\theta),\theta) - L - p\hat{q}(\theta)[1+\eta+\eta(\lambda-1)F(\theta)] \\ &\quad +\eta p\int_{\theta}^{\bar{\theta}}\hat{q}(\phi)f(\phi)d\phi + \eta\lambda p\int_{\underline{\theta}}^{\theta}\hat{q}(\phi)f(\phi)d\phi\,. \end{aligned} \tag{A.42}$$

Taking the derivative of $V(\cdot;\theta)$ with respect to λ yields

$$\begin{aligned} V'(\lambda;\theta) = \frac{d\hat{q}(\theta)}{d\lambda}\underbrace{\left[\frac{\partial u(\hat{q}(\theta),\theta)}{\partial q} - p[1+\eta+\eta(\lambda-1)F(\theta)]\right]}_{=0} - \eta pF(\theta)\hat{q}(\theta) \\ +\,\eta p\int_{\theta}^{\bar{\theta}}\frac{d\hat{q}(\phi)}{d\lambda}f(\phi)d\phi + \eta p\int_{\underline{\theta}}^{\theta}\hat{q}(\phi)f(\phi)d\phi + \eta\lambda p\int_{\underline{\theta}}^{\theta}\frac{d\hat{q}(\phi)}{d\lambda}f(\phi)d\phi. \end{aligned} \tag{A.43}$$

Note that

$$\frac{d\hat{q}(\lambda;\theta)}{d\lambda} = \frac{\eta pF(\theta)}{\partial^2 u(\hat{q}(\theta),\theta)/\partial q^2} \leq 0\,. \tag{A.44}$$

Thus, $V'(\lambda;\theta) \leq 0$ if

$$\hat{q}(\theta)F(\theta) - \int_{\underline{\theta}}^{\theta}\hat{q}(\phi)f(\phi)d\phi \geq 0, \tag{A.45}$$

which is satisfied, because $\hat{q}(\cdot)$ is non-decreasing.

The consumer's expected utility is given by $\mathbb{E}_\theta[V(\lambda;\theta)] = \int_{\underline{\theta}}^{\bar{\theta}} V(\lambda;\theta)f(\theta)d\theta$. Hence, the change in expected utility due to an increase in the consumer's degree of loss aversion is given by

$$\frac{d}{d\lambda}\mathbb{E}_\theta[V(\lambda;\theta)] = \int_{\underline{\theta}}^{\bar{\theta}} V'(\lambda;\theta)f(\theta)d\theta \leq 0.$$

Q.E.D.

Proof of Proposition 4: Irrespectively of the rival's tariff offer, if the sorting constraint is satisfied it is optimal for a firm to choose p_j such that $S_j(p_j)$ is maximized. Put differently, the firm will choose the method of generating v_j that maximizes its (per customer) profits. Thus, if no type $\lambda \in \{\lambda_1,\lambda_2\}$ has an incentive to mimic the other type, it is an equilibrium that the firms offer the same tariffs as in the full information case. Obviously, in case (ii) where $c \leq \Sigma(\lambda_1) < \Sigma(\lambda_2)$, both firms offer a flat-rate tariff to consumers. In this case, a flat-rate tariff maximizes $S_1(p)$ as well as $S_2(p)$. Moreover, the generated joint surplus is the same for both types of loss averse consumers. Since the brand preferences are i.i.d. across the λ_1 and λ_2 types, in any equilibrium each firm offers a single flat-rate tariff to consumers.

In the remaining part of the proof we show hat in the case where $\Sigma(\lambda_1) < c \leq \Sigma(\lambda_2)$ neither type λ_1 has an incentive to choose the tariff $(0, L^F)$ nor does type λ_2 have an incentive to choose the tariff $(\hat{p},\hat{L})$.

Claim $\hat{v}_1 \geq \hat{v}_2$.

Proof: Let $S_j^* \equiv \max_p\{S_j(p)\}$. Note that $S_1(0) = S_2(0) = S_2^*$. The firm's per customer profit from type $j = 1,2$ when offering utility v is

$$\pi_j(v) = S_j^* - v\,. \tag{A.46}$$

Thus, for any v it holds that $\pi_1(v) \geq \pi_2(v)$, since $S_1^* - v \geq S_2^* - v$. The equilibrium utilities are characterized by $\pi_j(\hat{v}_j) = \Phi(\hat{v}_j)$. Hence, we obtain the following relations:

$$\Phi(\hat{v}_1) = \quad \pi_1(\hat{v}_1) \geq \pi_2(\hat{v}_1) \tag{A.47}$$

$$\pi_1(\hat{v}_2) \geq \pi_2(\hat{v}_2) \quad = \Phi(\hat{v}_2)\,. \tag{A.48}$$

Suppose, in contradiction, $\hat{v}_1 < \hat{v}_2$. This immediately implies that $\pi_j(\hat{v}_1) > \pi_j(\hat{v}_2)$. Hence,

$$\Phi(\hat{v}_1) = \pi_1(\hat{v}_1) > \pi_1(\hat{v}_2) \geq \pi_2(\hat{v}_2) = \Phi(\hat{v}_2)\,. \tag{A.49}$$

Since $\Phi'(v) \geq 0$ the above formula holds only if $\hat{v}_1 > \hat{v}_2$, a contradiction. Q.E.D.

Since $\hat{v}_1 \geq \hat{v}_2$ and the expected utility from a flat-rate tariff being independent of λ, one can conclude that a consumer of type λ_1 has no incentive to choose the tariff $(0, L^F)$ that is designed for consumers of type λ_2. Finally, we show that type λ_2 has no incentive to mimic type λ_1. Let v_2^{DEV} denote the expected utility of a consumer of type λ_2 who accepts the tariff $(\hat{p},\hat{L})$ designed for type λ_1.

Claim $v_2^{DEV} < \hat{v}_2$.

Proof: The expected utility of type λ_2 from the tariff $(\hat{p}, \hat{L})$ equals the generated joint surplus minus the profits of the firm he purchases from. Thus,

$$v_2^{DEV} = S_2(\hat{p}) - \hat{L} - (\hat{p} - c) \int_{\underline{\theta}}^{\bar{\theta}} \hat{q}_2(\theta; \hat{p}) f(\theta) d\theta, \tag{A.50}$$

where $\hat{q}_2(\theta; p)$ denotes the demand of type λ_2 in the personal equilibrium. Inserting the explicit formula of $\hat{L}$ into (A.50) yields

$$v_2^{DEV} = S_2(\hat{p}) - \Phi(\hat{v}_1) - (c - \hat{p}) \int_{\underline{\theta}}^{\bar{\theta}} [\hat{q}_1(\theta; \hat{p}) - \hat{q}_2(\theta; \hat{p})] f(\theta) d\theta. \tag{A.51}$$

Note that $\hat{q}_1(\theta, \hat{p}) > \hat{q}_2(\theta, \hat{p})$ for all $\theta \in \Theta$, since $d\hat{q}/d\lambda < 0$ if $p > 0$. By Proposition 3 $c > \hat{p}$, and hence

$$v_2^{DEV} < S_2(\hat{p}) - \Phi(\hat{v}_1). \tag{A.52}$$

The expected utility of a consumer of type λ_2 when choosing the tariff that is designed for him can be expressed as follows,

$$\hat{v}_2 = S_2^* - \Phi(\hat{v}_2). \tag{A.53}$$

Hence, a deviation is not utility improving if

$$S_2^* - \Phi(\hat{v}_2) \geq S_2(\hat{p}) - \Phi(\hat{v}_1) \tag{A.54}$$

$$\iff [S_2^* - S_2(\hat{p})] + [\Phi(\hat{v}_1) - \Phi(\hat{v}_2)] \geq 0 \; . \tag{A.55}$$

The above inequality is satisfied since $\Phi'(\cdot) \geq 0$ and $\hat{v}_1 \geq \hat{v}_2$. Q.E.D.

Thus, if the firms offer the optimal tariffs of the full information case, each type of loss averse consumer selects the tariff that is designed for him, which completes the proof.

Q.E.D.

B. Examples of Discrete Choice Models

Hotelling Model with Linear Transport Cost.—Suppose consumers' ideal brands are uniformly distributed on the unit interval $[0, 1]$. The brands of the two firms, A and B, are located at the two extreme points, brand A at zero and brand B at one. A consumer with ideal brand $x \in [0, 1]$ has brand preferences $\zeta = (0, -tx, -t(1 - x))$. The parameter $t > 0$ is a consumer's "transport cost" per unit distance between his ideal brand and the brand he purchases from. For the Hotelling specification the market share function takes the following form,

$$m(v^A, v^B) = \min \left\{ \frac{1}{2t}(t + v^A - v^B), \frac{v^A}{t} \right\}. \tag{B.1}$$

The market share function has to be modified if v^A and v^B differ by so much that $m(\cdot) \notin [0, 1]$ (this never happens in equilibrium). Moreover, the Hotelling model has the well-known drawback that market shares are kinked. If, however, the transport cost is sufficiently low then one can focus on the case where the market share function is given by the first term of the above expression and thus well behaved. Formally, for $t \leq (2/3)S_2^*$ it suffices to analyze firms' profit maximization problem for[33]

$$m(v^A, v^B) = [1/(2t)](t + v^A - v^B). \tag{B.2}$$

Hence, $\partial m(v^A, v^B)/\partial v^A = (2t)^{-1}$ which immediately implies that

$$\Phi(v) \equiv \frac{m(v, v)}{\partial m(v, v)/\partial v^A} = t. \tag{B.3}$$

Obviously, $\Phi(\cdot)$ is non-decreasing. Note that

$$\frac{\partial m(v^A, v^B)/\partial v^A}{m(v^A, v^B)} = (t + v^A - v^B)^{-1}. \tag{B.4}$$

It can easily be seen that the above fraction is increasing in v^B. Thus, the Hotelling model satisfies all imposed assumptions if the transport cost is sufficiently low. One can check that the collusive utility level exists. To calculate the collusive utility level one has to use the market share function given in (B.1).

Logit Demand Model.—An obvious drawback of the Hotelling specification is that a firm does not compete with the rival and the outside option at the same time. A model that accounts for this simultaneous competition

[33]See Lemma 1 of Armstrong and Vickers (2001).

on two fronts is the logit demand model. Here, a consumer's brand preferences ζ^i for $i = 0, A, B$ are i.i.d. according to the double exponential distribution with mean zero and variance $\mu^2\pi^2/6$, where π (here) denotes the circular constant. Thus, the cumulative distribution function is

$$G(\zeta^i) = \exp\{-\exp[-(\gamma + \zeta^i/\mu)]\}, \tag{B.5}$$

where γ is the Euler–Mascheroni constant and μ is a positive constant. With this specification, the market share of firm A is given by (see Anderson et al. 1992)

$$m(v^A, v^B) = \frac{\exp[v^A/\mu]}{\exp[v^A/\mu] + \exp[v^B/\mu] + 1}. \tag{B.6}$$

The parameter μ captures the degree of heterogeneity among consumers with respect to their brand preferences. Put differently, μ measures the degree of product differentiation. A lower value of μ corresponds to a more competitive market. For $\mu \to \infty$ the firms are local monopolists. Taking the partial derivative of (B.6) with respect to v^A yields

$$\frac{\partial m(v^A, v^B)}{\partial v^A} = \frac{\exp[v^A/\mu]\{\exp[v^B/\mu] + 1\}}{\mu\{\exp[v^A/\mu] + \exp[v^B/\mu] + 1\}^2}. \tag{B.7}$$

Thus,

$$\frac{m(v^A, v^B)}{\partial m(v^A, v^B)/\partial v^A} = \frac{\mu\{\exp[v^A/\mu] + \exp[v^B/\mu] + 1\}}{\exp[v^B/\mu] + 1}. \tag{B.8}$$

Evaluating the above expression at $v^A = v^B = v$ leads to

$$\Phi(v) = \mu\frac{2\exp[v/\mu] + 1}{\exp[v/\mu] + 1}. \tag{B.9}$$

Taking the derivative of $\Phi(\cdot)$ with respect to v yields

$$\Phi'(v) = \frac{\exp[v/\mu]}{(\exp[v/\mu] + 1)^2} > 0. \tag{B.10}$$

Moreover, the derivative of $[\partial m(v^A, v^B)/\partial v^A][m(v^A, v^B)]^{-1}$ with respect to v^B amounts to

$$\frac{d}{dv^B}\left[\frac{\partial m(v^A, v^B)/\partial v^A}{m(v^A, v^B)}\right] = \frac{1}{\mu^2}\,\frac{\exp[v^B/\mu]\{\exp[v^B/\mu] + 1\}}{\mu\{\exp[v^B/\mu] + \exp[v^B/\mu] + 1\}^2} > 0. \tag{B.11}$$

The collusive utility level $\tilde{v}$ maximizes $m(v, v)\pi(v)$. Note that $m(v, v) \to 0$ for $v \to -\infty$ and $\pi(v) \leq 0$ if $v \geq \max_p\{S(p)\}$. Thus, the collusive utility exists, since $m(v, v)\pi(v)$ is continuously differentiable.

REFERENCES

[1] **Abeler, J., A. Falk, L. Götte,** and **D. Huffman (forthcoming):** Reference Points and Effort Provision, *American Economic Review*.

[2] **Anderson, S.P.**, **A. dePalma** and **J.-F. Thisse (1992):** Discrete Choice Theory of Product Differentiation, *MIT Press*, Cambridge.

[3] **Armstrong, M.** and **J. Vickers (2001):** Competitive Price Discrimination, *RAND Journal of Economics*, Vol. 32(4), 579-605.

[4] **Clay K.B.**, **D.S. Sibley** and **P. Srinagesh (1992):** Ex Post vs. Ex Ante Pricing: Optimal Calling Plans and Tapered Tariffs, *Journal of Regulatory Economics*, Vol. 4, 115-138.

[5] **Courty, P.** and **H. Li (2000):** Sequential Screening, *Review of Economic Studies*, Vol. 67, 697-717.

[6] **Crawford, V. P.**, and **J. Meng (2009):** New York City Cabdrivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income, http://dss.ucsd.edu/~vcrawfor/cabdriver11.3.pdf.

[7] **DellaVigna, S.** and **U. Malmendier (2004):** Contract Design and Self-Control: Theory and Evidence, *Quarterly Journal of Economics*, Vol. 119(2), 353-402.

[8] **DellaVigna, S.** and **U. Malmendier (2006):** Paying Not to Go to the Gym, *American Economic Review*, Vol. 96(3), 694-719.

[9] **Eliaz, K.** and **R. Spiegler (2006):** Contracting with Diversely Naive Agents, *Review of Economic Studies*, Vol. 73, 689–714.

[10] **Eliaz, K.** and **R. Spiegler (2008):** Consumer Optimism and Price Discrimination, *Theoretical Economics*, Vol. 3(4), 459–497.

[11] **Ericson, K.M.M.** and **A. Fuster (2009):** Expectations as Endowments: Reference-Dependent Preferences and Exchange Behavior, *Working Paper*, Harvard University.

[12] **Esteban, S.** and **E. Miyagawa (2006):** Temptation, Self-Control, and Competitive Nonlinear Pricing, *Economics Letters*, Vol. 90, 348–355.

[13] **Esteban, S.**, **E. Miyagawa** and **M. Shum (2007):** Nonlinear Pricing with Self-Control Preferences, *Journal of Economic Theory*, Vol. 135, 306-338.

[14] **Faulhaber, G.R.** and **C. Hogendorn (2000):** The Market Structure of Broadband Telecommunications, *Journal of Industrial Economics*, Vol. 68(3), 305-329.

[15] **Gerpott T.J. (2009):** Biased Choice of a Mobile Telephony Tariff Type: Exploring Usage Boundary Perceptions as a Cognitive Cause in Choosing Between a Used-Based or a Flat-Rate Plan, *Telematics and Informatics*, Vol. 26, 167-179.

[16] **Gill, D.** and **V. Prowse (2009):** A Structural Analysis of Disappointment Aversion in a Real Effort Competition, *IZA Discussion Paper*, No. 4536.

[17] **Goldman, M.B.**, **H.L. Leland** and **D.S. Sibley (1984):** Optimal Nonuniform Pricing, *Review of Economic Studies*, Vol. 51, 305-319.

[18] **Grubb, M. (2009):** Selling to Overconfident Consumers, *American Economic Review*, Vol. 99, 1770-1807.

[19] **Gul, F.** and **W. Pesendorfer (2001):** Temptation and Self Control, *Econometrica*, Vol. 60, 1293-1330. http://ssrn.com/abstract=717706.

[20] **Heidhues, P.** and **B. Kőszegi (2005):** The Impact of Consumer Loss Aversion on Pricing, *CEPR Discussion Paper*, No. 4849, http://ssrn.com/abstract=717706.

[21] **Heidhues, P.** and **B. Kőszegi (2008):** Competition and Price Variation when Consumers are Loss Averse, *American Economic Review*, Vol. 98, 1245–1268.

[22] **Heidhues, P.** and **B. Kőszegi (forthcoming):** Exploiting Naivete about Self-Control in the Credit Market, *American Economic Review*.

[23] **Herweg, F.**, **D. Müller** and **P. Weinschenk (forthcoming):** Binary Payment Schemes: Moral Hazard and Loss Aversion, *American Economic Review*.

[24] **Hobson, M.** and **R.H. Spady (1988):** The Demand for Local Telephone Service Under Optimal Local Measured Service, *Bellcore Economics Discussion Paper*, No. 50.

[25] **Karle, H.** and **M. Peitz (2010a):** Pricing and Information Disclosure in Markets with Loss-Averse Consumers, *CEPR Discussion Paper*, No. 7785.

[26] **Karle, H.** and **M. Peitz (2010b):** Consumer Loss Aversion and the Intensity of Competition, *Working Paper*, Université Libre de Bruxelles.

[27] **Kling, J.P.** and **S.S. van der Ploeg (1990):** Estimating Local Elasticities with a Model of Stochastic Class of Service and Usage, in *Telecommunications Demand Modeling: An Integrated View*, A. de Fontenay, M.H. Shugard and D.S. Sibley, editors, Amsterdam, North Holland, 119-136.

[28] **Kőszegi, B.** and **M. Rabin (2006):** A Model of Reference-Dependent Preferences, *Quarterly Journal of Economics*, Vol. 121, 1133-1165.

[29] **Kőszegi, B.** and **M. Rabin (2007):** Reference-Dependent Risk Preferences, *American Economic Review*, Vol. 97, 1047-1073.

[30] **Kridel, D.J.**, **D.E. Lehman** and **D.L. Weisman (1993):** Option Value, Telecommunication Demand, and Policy, *Information Economics and Policy*, Vol. 5 (2), 125-144.

[31] **Lambrecht, A.** and **B. Skiera (2006):** Paying Too Much and Being Happy About It: Existence, Causes, and Consequences of Tariff-Choice Biases, *Journal of Marketing Research*, Vol.63 (3), 212-223.

[32] **Lambrecht, A.**, **K. Seim** and **B. Skiera (2007):** Does Uncertainty Matter? Consumer Behavior Under Three-Part Tariffs, *Marketing Science*, Vol.26 (5), 698-710.

[33] **Leland, H.L.** and **R.A. Meyer (1976):** Monopoly Pricing Structures with Imperfect Discrimination, *Bell Journal of Economics*, Vol. 7, 449-462.

[34] **Macera R. (2009):** Intertemporal Incentives under Loss Aversion, *Working Paper*, University of California, Berkeley.

[35] **Maier, J.**, and **M. Rüger (2009):** Reference-Dependent Risk Preferences of Higher Orders, *mimeo*, University of Munich.

[36] **Maskin, E.** and **J. Riley (1984):** Monopoly with Incomplete Information, *RAND Journal of Economics*, Vol. 15, 171-196.

[37] **Miravete, E.J. (2002):** Estimating Demand for Local Telephone Service with Asymmetric Information and Optional Calling Plans, *Review of Economic Studies*, Vol. 69, 943-971.

[38] **Miravete E.J. (2003):** Choosing the Wrong Calling Plan? Ignorance and Learning, *American Economic Review*, Vol. 93 (1), 297-310.

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

[40] **Mitomo, H.**, **T. Otsuka** and **K. Nakaba (2009):** A Behavioral Economic Interpretation of the Preferences for Flat Rates: A Case of Post-Paid Mobile Phone Services, *mimeo*, Waseda University.

[41] **Mussa, M.** and **S. Rosen (1978):** Monopoly and Product Quality, *Journal of Economic Theory*, Vol.18, 301-317.

[42] **Nunes, J.C. (2000):** A Cognitive Model of People's Usage Estimations, *Journal of Marketing Research*, Vol. 37, 397-409.

[43] **Oi, W.Y. (1971):** A Disneyland Dilemma: Two-Part Tariffs for a Mickey Mouse Monopoly, *Quarterly Journal of Economics*, Vol. 85, 77-96.

[44] **Post, T., M.J. Van den Assem, G. Baltussen**, and **R.H. Thaler (2008):** Deal Or No Deal? Decision Making Under Risk in a Large-Payoff Game Show, *American Economic Review*, 98(1): 38-71.

[45] **Spiegler, R. (2010):** Monopoly Pricing when Consumers are Antagonized by Unexpected Price Increases: A "Cover Version" of the Heidhues-Kőszegi-Rabin Model, *Manuscript*, London School of Economics.

[46] **Sundararajan, A. (2004):** Nonlinear Pricing of Information Goods, *Marketing Science*, Vol.50 (12), 1660-1673.

[47] **Thaler, R.H. (1999):** Mental Accounting Matters, *Journal of Behavioral Decision Making*, Vol. 12, 183-206.

[48] **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 Choice, *RAND Journal of Economics*, Vol.18 (1), 109-123.

[49] **Train, K.E.**, **M. Ben-Akiva** and **T. Atherton (1989):** Consumption Patterns and Self-Selecting Tariffs, *Review of Economics and Statistics*, Vol.1, 62-73.

[50] **Train, K.E. (1991):** Optimal Regulation: The Economic Theory of Natural Monopoly, *MIT Press*, Cambridge.

[51] **Uthemann, A. (2005):** Competitive Screening of Customers with Non-Common Priors, *Working Paper*, University College London.

[52] **Willig, R.D. (1978):** Pareto-Superior Nonlinear Outlay Schedules, *Bell Journal of Economics*, Vol. 9, 56-69.

[53] **Wilson, R.B. (1993):** Nonlinear Pricing, *Oxford University Press*, New York.