Service Refusal in Regulated Markets for Credence Goods

Maarten C.W. Janssen* and Alexei Parakhonyak †

October 15, 2011

Abstract

This paper analyzes dynamic selection effects in markets for credence goods where price structures are determined by a regulator or by central management.There are different types of consumers and each type requires a different service or treatment level. We show that for a large class of price structures some types of consumers are not treated and refused the service. Equilibria with selection are welfare inferior to equilibria without selection. We also characterize the class of price structures for which selection does not arise. As the market becomes larger or service providers become more patient the class of selection-free price structures shrinks and in the limit it is unique. We show that this unique price structure also removes incentives for overtreatment.

JEL-codes: I11, L51

Key Words: Credence goods, Overtreatment, Selection effects

*Dept. of Economics, University of Vienna and Higher School of Economics, Moscow. Address: Hohenstaufengase 9, 1010 Vienna, Austria. Email: maarten.janssenunivie.ac.at.

†Higher School of Economics, corresponding author, Pokrovsky Boulevard, 11, 101000 Moscow, Russia. Phone: +7 (495) 772-95-90. Fax: +7 (495) 772-95-90. E-mail: parakhonyak@yandex.ru.

Earlier versions of this paper were circulated under the title “Dynamic Selection Effects in Regulated Markets” and were presented at Tinbergen Institute Rotterdam and at the European Meetings of the Econometric Society. We thank Josse Delfgaauw, Uwe Dulleck, Winand Emons and the audiences for helpful comments.

1 Introduction

Consider yourself arriving after a long trip at the railway station of your final destination. You know it is not too far to your hotel, but you want to take a taxi because of your luggage and because of the fatigue. Taxis are standing in line and have regulated non-negotiable fees. You walk up to the first taxi waiting in line and after hearing where you want to be taken, the driver tells you that you better walk because he refuses to take on passengers for such a short distance. As an economist, you may wonder: is this rational behaviour on the part of the taxi driver? If so, what is the role of the fare structure and does a fare structure exists where potential passengers are not refused? Can it be socially optimal that potential passengers are refused?

The taxi market is one of the classic examples of a market for credence goods (or the market for expert services) as first introduced in Darbi and Karni (1973). The main feature of credence goods is that “consumers do not know what kind of service they need, but they observe the utility of what they get” (Dulleck et al. (2011)). In the taxi example, the driver may in principle take a longer road, or apply a different fare. Medical doctors may provide unnecessary treatment (see, e.g. McGuire (2000)) and car mechanics can do repairs which are not needed, or charge for repairs that have not been done.

The taxi example presented in the first paragraph is interesting in that the taxi driver apparently felt he could or did not want to cheat (as he refused to provide the service instead of taking a longer road (provide overtreatment)). The fact that the service is not provided is an inefficiency: if you were able to negotiate the price, you probably would reach a mutually beneficial deal with the driver. For the price reflected on the taxi meter, however, the driver prefers to wait for a next client on whom he probably can make more money.

The problem of service refusal (or selection of clients) has not been the focus of attention in the literature on credence goods or the market for expert advice. The fundamental reason why service refusal may occur is similar, however, to the problem of mistreatment. This problem essentially says that the service provider treats one type of consumer as another type, and that servicing this other type is more profitable from the provider’s perspective. In case of service refusal there is an intertemporal substitution: the provider expects the next client needs a service that results in higher profits.

There are three aspects to the taxi example that are crucial for intertemporal selection to arise. First, different types of consumers need different service levels and they arrive sequentially at the service provider. In the taxi example: different passengers have different travel destinations and thus require different travel times. Second, the price structure (how price depends on type) is fixed by a regulatory agency or by central company management. In the taxi example: in different countries around the world, taxi drivers are not free to determine their own fare structure, but the fare structure is centrally programmed in the taxi meter. Third, agents who actually provide the service can either accept or reject customers based on a comparison of benefits and costs. In the taxi example: taxi drivers are “free” to tell potential clients that they do not take them or at least that it is very difficult to enforce a system where taxi drivers have to take all passengers. It is optimal for a taxi driver to refuse a passenger if the expected discounted revenue of waiting for the next passenger (the chance of getting a " big fish") is larger than the revenue of taking the current passenger and behave optimally thereafter.

The taxi market is not the only market where these features are present. In many countries, many parts of the medical sector also satisfy the main features outlined above. First, patients demanding some treatment enter a hospital or private clinic sequentially. Second, medical doctors are not free to set their own fees, but instead the fees per treatment are set by government authorities. Finally, medical doctors can refuse to take on patients and send them to other hospitals, sometimes giving the argument that other doctors are better equipped to provide the proper treatment. Other markets that have features that are described above include the market for social attorneys,[1] and some repair markets (car mechanics, shoes, electronics) where prices for standard repairs are set at the central management level and franchise holders bear the revenues and costs.

In this paper we analyze markets that are characterized by the three features mentioned above. We show that for a large class of price schedules, selection is a crucial aspect of the equilibrium in these markets: depending on the price schedule some types of consumers are refused (or treated improperly). We then characterize the (class of) price structures for which selection does not arise. These price structures can be set *a priori*, i.e., the regulator or central management does not have to observe the type of a particular customer. As the number of consumers increases or agents become more patient this class of selection-free price structures shrinks and in the limit it is unique. We also show that selection is always bad from a welfare point of view in the sense that for any price structure that gives rise to selection, there exists another price structure without selection that generates a higher total surplus.

Our framework can also be used to study demand inducement (or overtreatment). Medical doctors seem to provide more or a different treatment than what would be socially optimal for a patient with a particular disease.[2] Emons (1997) cites a Swiss study reporting that the average person's probability of receiving one of seven major surgical interventions is one third above that of a physician or a member of a physician's family. Overtreatment is also a common phenomenon in repair markets. Wolinsky (1993) cites a survey conducted by the U.S. Department of Transportation estimating more that 50% of car repairs being unnecessary. Balafoutas et al. (2011) conducted a natural field experiment with taxi drivers in Athens. They found that on average the length of a ride was 8% longer than the shortest possible road, and if a person is perceived

[1] In quite a few countries, low income families can apply for legal aid (attorneys) at a regulated fee. The fee structure has been modified recently in The Netherlands and this has lead to selection effects as attorneys argued the fee structure is such that they cannot provide legal aid to certain clients.

[2] The literature on demand inducement in health care markets started off with Evans (1974) and includes both theoretical (McGuire and Pauly (1991), Gruber and Owings (1996), Ellis (1998) and Wright (2007), among others) and empirical (such as Ellis and McGuire (1995)) papers.

to be a high-income foreigner it is 11% longer.

We show that the unique price structure that solves the inefficiency related to service refusal also solves the moral hazard problem of overtreatment.

Our paper is obviously related to the large literature on markets for credence goods, or the market for expert advice. For a relatively recent overview of this literature, we refer the reader to Dulleck and Kerschbamer (2006). Our paper is, in particular, related to two papers by Pitchik and Schotter (1987, 1993). These papers also use a framework where prices are exogenously given and ask (among other things) the question how the price structure affects the incentives to behave honestly. We pose this question in a dynamic, intertemporal environment that is not considered in Pitchik and Schotter (1987, 1993). In particular, our results suggest that by choosing a particular price structure, service providers may have no incentives to cheat. Another paper that is of special interest is Wolinsky (1995), where setting pricing oneself out of a submarket (which can be regarded as a special type of service refusal) is an essential ingredient to get experts specialize in certain treatments. In our model, experts do not have the choice of specializing, however, and prices are not set at the level of the service provider. The result in Fong (2005) that experts do cheat selectively can also arise in our paper for prices that are set optimally for some treatments, but not for others.

Our paper also makes some contribution to the literature on price structures in taxi markets. Most taxi markets around the globe are heavily regulated. Cities as diverse in nature as New York, London, Tokyo, Amsterdam, Shanghai and Singapore all have price structures which are regulated (see, e.g., Yang et al. (2004)). Although every city has a structure with an initial charge and a distance-based charge, the precise nature of the price structure differs from city to city.[3] What is striking is that certainly for smaller distances, almost all price structures are linear in distance. In some cities the price proportional to distance becomes lower, if distance is beyond a certain threshold. In this paper we show that such price structures always lead to selection of some customers. Glazer and Hassin (1983) were the first to analyze selection and cheating in the taxi market and derived cheat-proof prices. Their results are obtained, however, in a specific context with two types of customers and specific assumptions on the relation between the processing time and the equilibrium level of selection.

The rest of the paper is structured as follows. The next section describes the model in more detail. Section 3 presents our main results on the refusal-free price structures. Section 4 discusses the welfare aspects and section 5 extends the analysis to study overtreatment effects where suppliers have the possibility of giving consumers a treatment that is different from the required treatment. Section 6 concludes. All the proofs can be found in Appendix A.

[3]In addition, the fare may depend on delay-based charges, and additional week-end or night charges.

2 General Model

Consider a market where in each certain time interval Δt a consumer arrives and demands some service.[4] A consumer is of a certain type θ, where θ is randomly distributed over the interval $[\theta_{\min}, \theta_{\max}]$ according to some continuously differentiable distribution function $F(\theta)$. A consumer of type θ derives a utility $u(\theta)$ when he gets the service that he requires. The utility function $u(\theta)$ is continuous and the consumer's reservation utility is normalized to 0. The market is (centrally) regulated, which means that the price structure per unit $g(\theta)$ is fixed by a central authority or by central management. Note, that the central authority does not observe the type θ of a particular consumer; it just chooses a complete price structure for any θ.

There are N agents (service providers) who provide the service and in the basic model they simply decide whether or not to provide the required service to the consumer. Demand inducement or overtreatment is discussed in Section 4. The decision whether or not to provide the service takes place after the service provider is informed about the type θ of the consumer. When a service provider does not provide the service to the first consumer, he waits until the next consumer arrives. For simplicity, service providers have infinite planning horizons and maximize the expected present value of future cash flows. Payments are made at the moment the service is provided. The material costs of providing the required service for type θ are given by $c(\theta)$ and the time it takes to provide the required service to a consumer of type θ is given by $t(\theta, \Theta)$, where Θ is the set of consumers not serviced in equilibrium (see the discussion below). The time cost implies that when a service is provided no other consumer can be serviced during the time period $t(\theta, \Theta)$. We define $f(\theta) = g(\theta) - c(\theta)$ as the net price. In order to avoid trivialities we assume that $f(\theta) > 0$ for any $\theta \in [\theta_{\min}, \theta_{\max}]$.

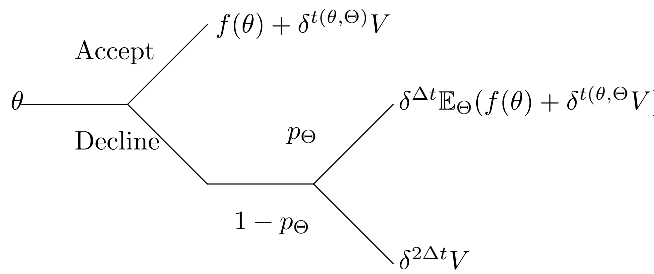
To understand why the servicing time $t(\theta, \Theta)$ depends on Θ, consider an equilibrium where some subset $\Theta \subset [\theta_{\min}, \theta_{\max}]$ of the consumers is not serviced. Let p_Θ be a probability that a service provided encounters an acceptable consumer: $p_\Theta = Pr(\theta \in [\theta_{\min}, \theta_{\max}] \setminus \Theta)$. The equilibrium level of consumers that is refused the service affects the processing time $t(\theta, \Theta)$. For example, to serve a consumer a taxi driver bears two types of time costs: he has to drive the consumer to his destination and he has to wait for a new consumer. Even though the former time cost almost entirely depends on the consumer's type (travel distance), the latter is mostly driven by the level of selection: namely on the proportion of free taxi cabs to the number of consumers.

The objective of the service providers is to maximize their discounted *expected* value of future cash flows *conditional on* θ: $v(\theta)$.[5] The *unconditional* expected value of future cash flows is denoted by V. It is important to em-

[4] This discrete time arrival model is not essential to the analysis. It can be shown that our results also hold in case consumers arrive according to a Poisson process. Details are available upon request.

[5] We abstract from discussions whether providers are also lead by different considerations, such as the health of their patients or the socially optimal level of service. We only say that if the provider is indifferent between providing any type of service, he will choose to provide the optimal service.

Figure 1: The decision tree of the service provider



phasize that the value of V, which is the value *prior to* the moment when a service provider knows the next consumer's type, should be distinguished from the value of future cash flows after the type is revealed. This latter value is $v(\theta)$ and can be either larger or smaller than V. The discount factor δ is assumed to be the same across all service providers. As the model horizon is infinite and as each time the service provider has to make a decision whether or not to service a consumer he faces the same situation, V is constant over time.

We will say that there is monotone selection, if the following is true: if a certain type $\widehat{\theta}$ is not serviced, then all types $\theta > \widehat{\theta}$ (or all types $\theta < \widehat{\theta}$) are also not serviced. In case of monotone selection, the agent's decision-making problem can be presented by the tree in Figure 1.

Once the service provider knows a consumer's type, he can either deliver or refuse to provide the required service. If he delivers, he receives the net price $f(\theta)$ and some expected continuation value. If he refuse to provide the service, an acceptable consumer arrives with probability p_Θ in the next period. The service provider compares the value of providing the required service to a consumer of type θ now, given by $f(\theta) + \delta^{t(\theta,\Theta)}V$, with the expected value of refusing to provide, which equals

$$p_\Theta \left[\delta^{\Delta t}\mathbb{E}_\Theta(f(\theta) + \delta^{t(\theta,\Theta)}V)\right] + (1 - p_\Theta)\,\delta^{2\Delta t}V.$$

where $\mathbb{E}_\Theta$ represents the expectation given the equilibrium level of selection Θ, i.e. the expectation over types $[\theta_{\min}, \theta_{\max}] \setminus \Theta$. If the pricing structure is such that every consumer is serviced, then the service provider should at least be weakly better off by servicing any type of consumer.

3 Analysis

The analysis proceeds by first considering the conditions that need to be satisfied for an equilibrium to exist where every consumer is serviced.

3.1 No Service Refusal

We first characterize the (class of) price structure(s) that is such that all consumers are serviced in equilibrium. Note that in this case $p_\Theta = 1$ so that the continuation value V is defined by:

$$V = \mathbb{E}(f(\theta) + \delta^{t(\theta,\emptyset)}V), \tag{3.1}$$

where $t(\theta,\emptyset)$ equals $t(\theta,\Theta)$ in case no types are refused the service they require. To satisfy the incentive constraint of the service provider we must have that for any θ in the support of $F(\theta)$ the following holds:

$$f(\theta) + \delta^{t(\theta,\emptyset)}V \geq \delta^{\Delta t}V \tag{3.2}$$

where V is defined above.

Thus, we can formulate the following proposition:

Proposition 3.1. *A price structure $f(\theta)$ insures no service refusal, if and only if, equations* (3.1),(3.2) *hold for $f(\theta)$.*

In general for a given price structure $f(\theta)$, service providers may still prefer one type above another even if there is no service refusal. But since Δt is finite, the service providers still do not have an incentive to refuse some consumers as it is still better to service the least profitable type θ now than to service an average type later.

A particular example of an optimal price structure arises when service providers are indifferent between servicing any type of consumer. If this is the case then even if Δt goes to 0, service refusal does not arise in equilibrium. We define this particular price structure in the following proposition.

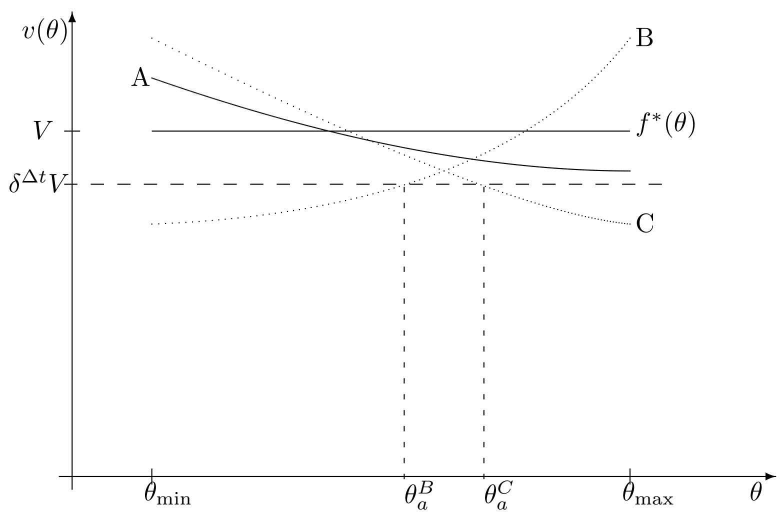
Proposition 3.2. *The price structure that makes service providers indifferent between servicing any type of consumer and thereby eliminating service refusal is defined by:*

$$f^*(\theta) = V(1 - \delta^{t(\theta,\emptyset)}) \tag{3.3}$$

where V is continuation value.

The price structure $f^*(\theta)$ achieves what is sometimes called an equal compensation principle (see, e.g., Milgrom and Roberts (1992), pp. 228-32). In the organization literature, this principle is invoked in situations where an employer cannot monitor an employee's allocation of attention between different activities. If an employee is expected to devote time to an activity for which the performance cannot be measured, then incentive pay cannot work (or in other words, the employee should be indifferent between allocating time to different activities). The price structure $f^*(\theta)$ implies that the prices have to reflect the material cost and the time spent on a treatment. Therefore, in general price structure $f^*(\theta)$ does not coincide with a fixed price independent of the type that has to be treated. Note that the value of V is not determined here. It is a free parameter determining the level of prices in the model and we can assign any arbitrary value to it.

Figure 2: Agent's pay-off under different price structures



We can summarize our results so far by means of Figure 2. The expectation of $v(\theta)$ for all prices must be equal to V. For the price structure $f^*(\theta)$ we have that $v(\theta)$ is a constant, i.e., $v(\theta) = V$. Price structure A does not imply a constant value: service providers prefer lower types of consumers. Given the waiting time Δt, service providers still find it optimal, however, to service all consumers. If the arrival frequency of consumers increases this price structure will induce that the service will be denied to relatively high types. Price structure B induces that low types will not be serviced. In particular, consumers of type θ smaller than θ_a^B are denied the service they require. The price structure C leads to an outcome where all consumers with a type larger than θ_a^C are not accepted.

3.2 Service Refusal at busy places

A natural question to ask is what will happen at busy places, where the time that elapses between consecutive consumers arriving is very small. We will show that in case Δt is arbitrarily small, all service refusal-free price structures are arbitrarily close to $f^*(\theta)$, i.e., any price structure other than the price structure where service providers are indifferent between servicing any type of consumer leads to service refusal.

We first introduce the class of service refusal-free price structures. Since the value V determines the level of prices, we consider structures with the same V.

Definition 3.3. *For any δ and Δt let the class of service refusal-free price* *structures, denoted by $\mathcal{F}_V(\delta, \Delta t)$, be defined by price structures $f(\theta)$ that satisfy the following three requirements:*

- for any $f(\theta) \in \mathcal{F}_V(\delta, \Delta t)$ there is no service refusal, i.e., condition (3.2) is satisfied for all θ in the support of the distribution;
- any $f(\theta) \in \mathcal{F}_V(\delta, \Delta t)$ gives the service provider the expected value V;
- $u(\theta) > f(\theta)$ for any $f(\theta) \in \mathcal{F}_V(\delta, \Delta t)$ and all θ.

If for all types the utility derived from a service is large enough, then certainly $f^*(\theta)$ is in $\mathcal{F}_V(\delta, \Delta t)$, and hence, $\mathcal{F}_V(\delta, \Delta t)$ is not empty. Using this definition, we can prove the following proposition.[6]

Proposition 3.4. *For any $\epsilon > 0$ there is a Δt^* such that for any $\Delta t^< \Delta t^*$ all service refusal-free price structures defined on $[\theta_{\min}, \theta_{\max}]$ are ϵ-close to the optimal one, i.e., they satisfy the following two conditions:*

1. $f^*(\theta) - f(\theta) < \epsilon$;
2. $\int_{\theta_{\min}}^{\theta_{\max}} |f^*(\theta) - f(\theta)| dF(\theta) < 2\epsilon(\theta_{\max} - \theta_{\min})$.

Intuitively, when Δt becomes arbitrarily small, the cost of waiting for the next consumer becomes small as well and waiting gives you the expected continuation pay-off. In this case, the only price structure that does not give rise to selection is the one where every type yields the same revenue (which is equal to the expected value). This is exactly how the price structure $f^*(\theta)$ is characterized.

4 Social welfare

Next, we investigate the welfare issues arising from selection. Our analysis shows that service refusal is never optimal from the point of view of social welfare in the sense that price structures that induce service providers to service all consumers are better than other price structures.

We define social welfare as the sum of consumer and producer (service provider) surplus. Producer surplus is simply equal to the discounted value of future pay-offs for the service providers multiplied by the number of them. Each consumer has a utility of $u(\theta) - f(\theta)$ per service provided. Integrating over all consumers that are serviced in equilibrium we arrive at the expected consumer surplus. Thus, social welfare conditional on V is given by:

$$SW(V) = N \cdot V + M \int_S (u(\theta) - f(\theta)) dF(\theta) \tag{4.1}$$

[6] A similar proposition can be proven for the case where the discount factor δ goes to 1. To economize on space, this proposition is not included in the text.

where $S = [\theta_{\min}, \theta_{\max}] \setminus \Theta$ is the support of the distribution when some types of consumers are denied the service and M is the relative weight that is put on consumer surplus. Integration over S_Φ implies that we calculate the surplus only over those types of consumers that are serviced in equilibrium.

Proposition 4.1. *Consider a price structure $g(\theta)$ which induces service refusal and delivers value V to the service providers. Then any $f(\theta) \in \mathcal{F}_V(\delta, \Delta t)$ yields a level of social welfare that is larger than $g(\theta)$.*

Whether or not the socially optimal price is necessarily also a Pareto-improvement depends on the time it takes to service different types of consumers. Indeed, for the optimal price structure $f^*(\theta)$ we have:

$$f^*(\theta) = (1 - \delta^{t(\theta,\emptyset)})V. \tag{4.2}$$

For a price structure with service refusal Θ we have:

$$g(\theta) = (\delta^{\Delta t} - \delta^{t(\theta,\Theta)})V. \tag{4.3}$$

Note, that if $t(\theta, \emptyset)$ is close to $t(\theta, \Theta)$ (e.g., when the time it takes to service a consumer is exogenous) consumers of type θ which are not in Θ, but sufficiently close to it are better off under $g(\theta)$. Thus, if each type of consumer always requires the same service, then the optimal price structure may not be Pareto-optimal. However, if consumers' types vary over time in the sense that a consumer requires different services over time, then consumers are interested in their expected surplus rather than in the surplus generated for a particular value of θ, and in that case the price structure that maximizes social welfare is also Pareto-improving.

5 Overtreatment

The framework developed so far also allows us to analyze the overtreatment (or moral hazard) problem mentioned in the Introduction. In order to do so we relax the assumption that either the right type of service has to be provided or no service at all. in particular, in this section we allow for any level of fraud: an allergy can be treated by abdominal surgery, a flat tire can be fixed by replacing also the engine, and it is possible for a taxi driver to make arbitrarily long detours. We show that the optimal price structure we propose removes the incentives to provide overtreatment even in this extreme case. If the possible extent of the fraud is limited, the results of this section are still applicable.

Assume the service provider can provide the consumer with any level of service. Thus, if the service provider decides to provide the service that is optimal for type θ_0, then the required treatment time is $t(\theta_0, \emptyset)$[7] and the payment to the service provider is given by the net price $f(\theta_0)$, independent of the true value

[7] When the choice of treatment is fully determined by the service provider, no consumer will be refused the service as the service provider will always choose the treatment that is optimal from his perspective.

of θ. The service provider can substitute the true θ with θ_0 and perform (and be paid for) the service that is required by θ_0.

It is clear that if there is the possibility to cheat on a consumer, the option of service refusal is no longer relevant as the service provider can always provide the service that is best *from his perspective*. We denote the type of this best consumer by $\theta_0 \in \text{Argmax}_\theta(f(\theta) + \delta^{t(\theta,\emptyset)}V)$. Then, the service provider can either service the consumer truthfully or cheat.

The next Proposition argues that under the optimal price structure, the service provider does not have an incentive to induce extra demand (cheat).

Proposition 5.1. *Under the optimal price structure 6.1 there is no pair (θ, θ_0) from the support of $F(\theta)$ such that θ_0 is better for the service providers than θ.*

Thus, the optimal price structure $f^*(\theta)$ we characterize allows to avoid both service refusal and moral hazard (demand inducement) problems. It does so by alluding to the equal compensation principle mentioned before that service providers should be made indifferent between providing the different services. We show it is possible to design such an optimal price structure that service providers indeed are indifferent between providing (and getting paid for) any possible treatment and so they do not have an incentive to cheat. If they have a slight (social) preference for doing good (providing the optimal, socially efficient, treatment) they will do so.

6 Conclusions

In this paper we have analyzed the effect of regulated price structures on the decision of service providers to deny services or to provide non-required services in markets for credence goods. Our framework applies when three core conditions are satisfied: (i) consumers differ in their types (service required) and arrive sequentially in time; (ii) price structures, with price depending on a consumer's type, are fixed by a regulator or by central management and (iii) service providers can freely decide on the service themselves and service truthfully, deny the service or cheat and give a different treatment.

We have shown that for a large class of fare structures, service refusal may arise in regulated markets for credence goods. Equilibria with service refusal are welfare inferior to equilibria where service refusal does not occur. We have also characterized the class of price structures for which service refusal does not arise. For large markets, this price structure is unique up to a scaling factor.

Bibliography

L. Balafoutas, A. Beck, Kerschbamer R., and M. Sutter. What drives taxi drivers? a field experiment on fraud in a market for credence goods. University of Innsbruck Working Papers in Economics and Statistics, 2011.

M. Darbi and E. Karni. Free competition and the optimal amount of fraud. *Journal of Law and Economics*, 16:67–88, 1973.

U. Dulleck and R. Kerschbamer. On doctors, mechanics, and computer specialists: The economics of credence goods. *Journal of Economic Literature*, 44: 5–42, 2006.

U. Dulleck, R. Kerschbamer, and M. Sutter. The economics of credence goods: An experiment on the role of liability, verifiability, reputation, and competition. *American Economic Review*, 101:526–555, 2011.

R.P. Ellis. Creaming, skimping and dumping: Provider competition on the intensive and extensive margins. *Journal of Health Economics*, 17(5):537–555, 1998.

R.P. Ellis and T.G. McGuire. Hospital response to prospective payment: Moral hazard, selection, and practice-style effects. *Journal of Health Economics*, 15 (3):537–555, 1995.

W. Emons. Credence goods and fraudulent experts. *RAND Journal of Economics*, 28:107–119, 1997.

R. Evans. Supplier-induced demand: Some empirical evidence and implications. In *The Economics of Health Care (ed. M. Perelman)*, pages 162–173. 1974.

Y.. Fong. When do experts cheat and whom do they target? *RAND Journal of Economics*, 36:113–130, 2005.

A. Glazer and R. Hassin. The economics of cheating in the taxi market. *Transportation Research Part A*, 17:25–31, 1983.

J. Gruber and M. Owings. Physician incentives and cesarean section delivery. *RAND Juornal of Economics*, 27:99–123, 1996.

T.G. McGuire. Chapter 9. In *Handbook of Health Economics*, pages 461–536. 2000.

T.G. McGuire and M.V. Pauly. hysicians response to fee changes with multiple payers. *Journal of Health Economics*, 10(4):385–410, 1991.

P. Milgrom and J. Roberts. *Economics, Organization and Management*. Prentice-Hall, New-York, 1992.

C. Pitchik and A. Schotter. Honesty in a model of strategic information. *American Economic Review*, 77:1032–1036, 1987.

C. Pitchik and A. Schotter. Information transmission in regulated markets. *Canadian Journal of Economics*, 26:815–829, 1993.

A. Wolinsky. Competition in a market for informed experts' services. *RAND Journal of Economics*, 24:380–398, 1993.

A. Wolinsky. Competition in markets credence goods. *Journal of Industrial and Theoretical Economics*, 151:117–131, 1995.

D.J. Wright. Specialist payment schemes and patient selection in private and public hospitals. *Journal of Health Economics*, 26(5):1014–1026, 2007.

H. M. Ye Yang, H. Tang, and S.C. Wong. Regulating taxi services in the presence of congestion externality. *Transportation Research Part A*, 39:17–40, 2004.

Appendix A: Proofs

Proposition 3.2. *The price structure that makes service providers indifferent between servicing any type of consumer and thereby eliminating service refusal is defined by:*

$$f^*(\theta) = V(1 - \delta^{t(\theta,\emptyset)}) \tag{6.1}$$

where V is continuation value.

Proof. Since service providers are indifferent between servicing any type of consumer the continuation value must be constant in θ and equal to the expected value V. Thus, the equation

$$V = f^*(\theta) + \delta^{t(\theta,\emptyset)}V \tag{6.2}$$

must hold for any θ as an identity. Therefore,

$$f^*(\theta) = V(1 - \delta^{t(\theta,\emptyset)}). \tag{6.3}$$

Now we need to check whether the conditions (3.1) and (3.2) hold. For (3.1) we obtain:

$$V = \mathbb{E}\left(V(1 - \delta^{t(\theta,\emptyset)}) + \delta^{t(\theta,\emptyset)}V\right) = V. \tag{6.4}$$

For (3.2) we obtain:

$$V(1 - \delta^{t(\theta,\emptyset)}) + \delta^{t(\theta,\emptyset)}V = V > \delta^{\Delta t}V. \tag{6.5}$$

Thus, both conditions are satisfied and the proof is complete.

□

Proposition 3.4. *For any $\epsilon > 0$ there is a Δt^* such that for any $\Delta t^< \Delta t^*$ all service refusal-free price structures defined on $[\theta_{\min}, \theta_{\max}]$ are ϵ-close to the optimal one, i.e., they satisfy the following two conditions:*

1. $f^*(\theta) - f(\theta) < \epsilon$;
2. $\int_{\theta_{\min}}^{\theta_{\max}} |f^*(\theta) - f(\theta)| dF(\theta) < 2\epsilon(\theta_{\max} - \theta_{\min})$.

Proof. 1. Recall, that by definition

$$f^*(\theta) + \delta^{t(\theta,\emptyset)}V = V$$

for all θ in the support of $F(\theta)$. Also note, that since $f(\theta) \in \mathcal{F}_V(\delta, \Delta t)$ we have that for all θ

$$f(\theta) + \delta^{t(\theta,\emptyset)}V \geq \delta^{\Delta t}V.$$

By taking difference we obtain

$$f^*(\theta) - f(\theta) \leq (1 - \delta^{\Delta t})V$$

then by choosing $\Delta t^* = \frac{\ln(1-\frac{\epsilon}{V})}{\ln \delta}$ for any $\Delta t < \Delta t^*$ we get that

$$f^*(\theta) - f(\theta) < \epsilon$$

which proves the first part of the proposition.

2. To prove the second part recall, that since both price structures belong to $\mathcal{F}_V(\delta, \Delta t)$ we have

$$\mathbb{E}f^*(\theta) + V\mathbb{E}\delta^{t(\theta,\emptyset)} = V$$
$$\mathbb{E}f(\theta) + V\mathbb{E}\delta^{t(\theta,\emptyset)} = V$$

and therefore

$$\int_{\theta_{\min}}^{\theta_{\max}} [f^*(\theta) - f(\theta)]dF(\theta) = 0 \tag{6.6}$$

Let A^+ be a set of all θ in the support of distribution, such that $f^*(\theta) \geq f(\theta)$, and A^- be a set such that $f^*(\theta) < f(\theta)$, $A^+ \bigcup A^- = [\theta_{\min}, \theta_{\max}]$. Then

$$\int_{\theta_{\min}}^{\theta_{\max}} |f^*(\theta) - f(\theta)|dF(\theta) = \int_{A^+} [f^*(\theta) - f(\theta)]dF(\theta) - \int_{A^-} [f^*(\theta) - f(\theta)]dF(\theta)$$

From (6.6) we get

$$0 = \int_{A^+} [f^*(\theta) - f(\theta)]dF(\theta) + \int_{A^-} [f^*(\theta) - f(\theta)]dF(\theta)$$

Therefore, by taking the difference and using the first part of the proposition we obtain:

$$\int_{\theta_{\min}}^{\theta_{\max}} |f^*(\theta) - f(\theta)| dF(\theta) = 2 \int_{A^+} [f^*(\theta) - f(\theta)] dF(\theta) < 2\epsilon(\theta_{\max} - \theta_{\min})$$

□

Proposition 4.1. *Consider a price structure $g(\theta)$ which induces service refusal and delivers value V to the service providers. Then any $f(\theta) \in \mathcal{F}_V(\delta, \Delta t)$ yields a level of social welfare that is larger than $g(\theta)$.*

Proof. Consider some $f(\theta) \in \mathcal{F}_V(\delta, \Delta t)$. Since both price structures $g(\theta)$ and $f(\theta)$ deliver value V to the service providers, producer surplus equals NV in both cases. On the other hand, under price structure $g(\theta)$ fewer consumers are serviced. Therefore to provide the same V they must pay on average more:

$$\int_S f(\theta) dF(\theta) < \int_S g(\theta) dF(\theta). \tag{6.7}$$

Let W be the consumer surplus. Then, using (6.7) we obtain:

$$\begin{aligned} W(f(\theta)) = \int_{\theta_{\min}}^{\theta_{\max}} (u(\theta) - f(\theta)) dF(\theta) > \int_S (u(\theta) - f(\theta)) dF(\theta) > \\ > \int_S (u(\theta) - g(\theta)) dF(\theta) = W(g(\theta)) \end{aligned}. \tag{6.8}$$

Since the surplus of the service providers is constant, it follows that for any positive value of M social welfare is larger in case all types are serviced. □

Proposition 5.1. *Under the optimal price structure 6.1 there is no pair (θ, θ_0) from the support of $F(\theta)$ such that θ_0 is better for the service providers than θ.*

Proof. Recall, that the optimal price structure is defined by

$$f^*(\theta) = V(1 - \delta^{t(\theta, \emptyset)}).$$

Then the expected value of providing the service for any θ after it has been observed is defined by:

$$f^*(\theta) + \delta^{t(\theta,\emptyset)} V = (1 - \delta^{t(\theta,\emptyset)}) V + \delta^{t(\theta,\emptyset)} V = V = f^*(\theta_0) + \delta^{t(\theta_0,\emptyset)} V, \tag{6.9}$$

which is independent of the type.

□