Adding Waiting Time Penalties to the Hotelling Model

Olivier Gallay*, Max-Olivier Hongler

Ecole Polytechnique Fédérale de Lausanne (EPFL)
STI-IPR-LPM
Station 17
CH-1015 Lausanne

Abstract

We study the market partition between two competing firms that deliver services to waiting-time sensitive customers. In our model, the incoming customers select a firm on the basis of its posted price, the expected waiting time and its brand. More specifically, we quantify by a cost any departure from the ideal brand expected by each incoming customer. Considering that the two underlying queueing processes operate under high traffic regimes, we analyze the market sharing dynamics by using a diffusion process. As a function of control parameters, such as the waiting and brand departure costs or the incoming traffic intensity, we are able to analytically characterize a transition between an Hotelling-like regime (dominated by brand considerations) and a deadline type regime (dominated by waiting time considerations). The market sharing dynamics is described by the time evolution of a boundary point, which time evolution belongs to the class of noise-induced phase transitions, so far widely discussed in physics, chemistry and biology.

Keywords: Stochastic processes, Hotelling model, heavy traffic queueing dynamics, multiplicative noise, noise-induced phase transition.

*Corresponding author. Tel: +41 21 6935817; fax: +41 21 6933891; e-mail address: olivier.gallay@epfl.ch

1 Introduction

In his original contribution [Hotelling 1929], H. Hotelling did consider the case where two vendors supply an identical product that is perceived homogenous by incoming customers. However, the vendors being separated in geographical space, transportation costs to be added to the mill prices charged by the vendors are generated. In presence of two vendors, it exists an inner market *boundary point*, for which the mill price plus the transportation costs from both suppliers break even. This seminal modeling framework has stimulated a wealth of contributions with the goal to relax some of the oversimplifying hypothesis of the original model. In particular, the introduction of *elastic demands* (i.e. customers are not prepared to pay "prohibitive prices" for the product) has been discussed in [Puu 2002]. Note that the original Hotelling's model is basically deterministic - it indeed does not incorporate random perturbations which actually may corrupt the prices and then affect the customers' decision process. Among the numerous potential noise sources, one of the simplest and most natural way to incorporate randomness is to consider the situations where the customers' decision to select one of the vendors depends on the expected time delay before service. This simple and realistic situation has been recently proposed by G. Cachon and P. Harker in [Cachon et al. 1999] and [Cachon et al. 2002]. As these authors clearly emphasized in [Cachon et al. 1999], the resulting inherent analytical complexity implies that rather seldom are the models dealing with firms that simultaneously compete with both prices and processing rates. The aim of this note is to investigate a class of simple models for which explicitly analytical considerations can partly be worked out. While in [Cachon et al. 1999] the firms are assumed to adjust their processing rates to guarantee a fixed expected time cost, our class of models takes into account the fluctuations of the waiting times and therefore keeps full track of the randomness induced by the underlying queueing processes. Note that the adjusting processing rates assumption proposed in [Cachon et al. 1999] allows a discussion based only on averages. Contrary to [Cachon et al. 1999], where no variance effects enter into the description of the model (i.e. this is effectively a "pseudo-stochastic" model), our approach explicitly emphasizes the role played by the fluctuations variance - also called in the sequel the "volatility" of the underlying noise sources. As discussed in [Hassin et al. 2003], the introduction of waiting costs in the queueing dynamics leads to the concept of *externalities* (i.e. the costs induced on later incomers by a customer who is just joining the queue). In the class of models to be discussed here, these externalities trigger the random dynamics controlling the boundary point which defines the market partition.

In section 2, we show that, for heavy traffic regimes of the underlying queueing processes, the boundary point partitioning the market interval is governed by a scalar stochastic differential equation with multiplicative noise source. For this dynamics, it is possible to explicitly calculate the associated stationary probability measure. The multiplicative character of the noise source offers the possibility to observe a so-called *noise-induced phase transition*, which manifests itself by a change of the modal character of the stationary probability measure - in the simplest case realized here, a transition from uni- to a bimodal probability density. In the present context, the transition between these two regimes relates to a transition between a regime where the Hotelling's spatial (i.e. the brand) character dominates in the decision taken by the incoming customers and a regime where the time delays dominate. In section 3, we explicitly work out a simple, though fully representative, example from our class of models. For this particular choice of the dynamics, we are also able to fully calculate the relaxation rate (i.e. the transient regimes) characterizing the approach towards the final statistical equilibrium. The relaxation process is strongly dependent on the relative importance of the externalities arising in the associated queueing processes. Finally, a short account devoted to simulation experiments explicitly comforts our analytical findings.

2 Model

As in [Cachon et al. 1999], our starting point will be a two servers Hotelling model where two service providers S_1 and S_2 are located in a (linear) market confined on a segment $\Omega := [-\Delta, +\Delta] \subset \mathbb{R}$, $\Delta > 0$. The positions of the service providers are respectively denoted by $-\Delta \leq x_1 \leq 0$ and $0 \geq x_2 \geq +\Delta$ and are assumed to be symmetric with respect to the center of the market, i.e. $x_1 = -x_2$ and $L = 2x_2$ denotes the distance between S_1 and S_2. The servers S_1 and S_2 offer homogenous services and both charge an equal price p. Departing now from the original Hotelling's model, we add queueing processes in front of S_1 and S_2 and following [Hassin et al. 2003], we will attach waiting costs to any customer lining in the queues before being served. Taking into account waiting costs thus confers a dynamical character to the original Hotelling's model. Specifically, our dynamic model exhibits the following features and obeys to the following rules:

a) *Arrivals dynamics.* Incoming customers follow a Poisson rule with rate Λ, hence the average time between two arrivals will be Λ^{-1}.

b) *Spatial distribution of the arrivals.* Incoming customers arrive at a random location $x \in \Omega$ drawn from a uniform probability density $U(\Omega)$ with support on Ω.

c) *Services dynamics.* Both servers S_i, $i = 1, 2$, have generally distributed service times with rate μ_i, hence the average service time will be μ_i^{-1}, $i = 1, 2$.

d) *Traffic intensity.* The traffic into the system is limited to $\rho := \frac{\Lambda}{\mu_1 + \mu_2} < 1$. This ensures that the system is globally stable, i.e. the global incoming rate is less than the global service rate. In the sequel, we shall assume that both servers share a common rate $\mu = \mu_1 = \mu_2$.

e) *Queueing processes.* When an incoming customer finds both S_1 and S_2 busy, he/she will wait for service and line-up in a queue. The capacity of the queue is assumed to be unlimited and the service discipline is first-in-first-out (FIFO). In view of points a) and c), we hence consider M/G/1 queues.

f) *Customer information gathering.* Upon his/her arrival at $x \in \Omega$, each incoming customer knows:

> 1) his/her relative distance $\mid x - x_1 \mid$ and $\mid x - x_2 \mid$ to S_1 and S_2.
>
> 2) the contents $N_1(t)$ and $N_2(t)$ of both queues ($t \in \mathbb{R}^+$ being the arrival time). In other words, both queue contents are observable to any incoming customer.

g) *Cost structures.* There are two types of costs incurred by any customer, namely:

> 1) the waiting time cost (WTC), characterized by a cost parameter c_w with physical unit $\left[\frac{\text{dollar}}{\text{time unit}}\right]$.
>
> 2) the brand departure cost (BDC), quantified by a cost parameter c_t with physical unit $\left[\frac{\text{dollar}}{\text{brand distance unit}}\right]$.

h) *Decision policy.* Upon arrival and based on information regarding:

- the observed queue lengths $N_1(t)$ and $N_2(t)$,
- his/her relative position to S_1 and S_2,
- the values of the costs c_w and c_t,

any incoming customer decides which server S_1 or S_2 he/she will join.

i) *Demand structure*. Following the original Hotelling case, we assume an inelastic demand, i.e. a customer will purchase the service at any price, even if the proposed price is arbitrarily large.

A graphical sketch of our modeling framework can be found in Fig. 1.

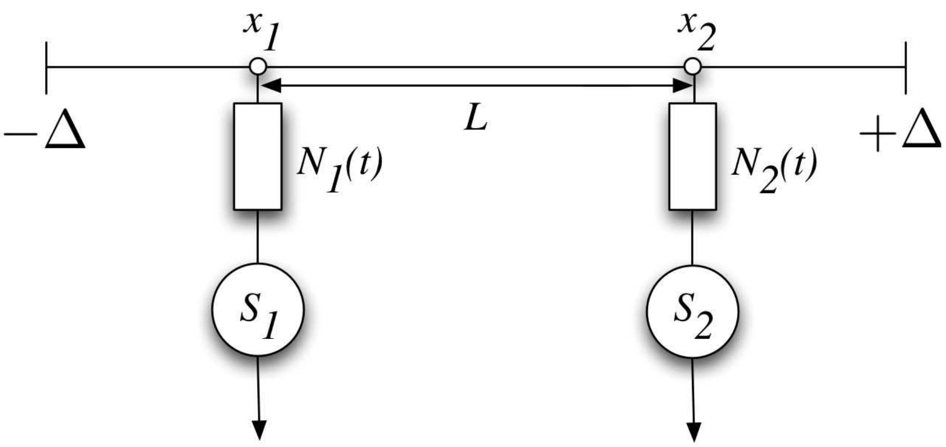

Figure 1: Bounded market with two vendors and time sensitive customers.

Different from the original Hotelling's idea, waiting times confer to the above class of models an explicit dynamic character.

When served by S_i, an incoming customer feels a utility function $U_i(x)$, $i = 1, 2$ and x is the customer's initial position which enters into the decision policy. In words, the functions $U_i(x)$ quantify the gain realized by a customer choosing server S_i when its entering location is x. Specifically, for linear waiting and transportation costs, the utility functions read as:

$$U_i(x) = a - p - c_t|x - x_i| - c_w \mathbb{E}\left(W_i|N_i(t)\right), \quad i = 1, 2, \tag{1}$$

with a being a systematic reward due to the service and $\mathbb{E}\left(W_i|N_i(t)\right)$ standing for the conditional expected waiting time at S_i when $N_i(t)$ already waiting customers are observed. As μ^{-1} is the average service time, this last conditional expectation is readily given by:

$$\mathbb{E}\left(W_i|N_i(t)\right) = \frac{N_i(t)}{\mu}.$$

We obviously assume that any customer does maximize his utility function when choosing one of the two servers. This suggests to introduce a *time-dependent boundary* position $Y_t \in [-\Delta, +\Delta]$ to be a separation point implicitly defined by:

$$U_1(Y_t) = U_2(Y_t). \tag{2}$$

Hence, our strictly increasing (BDC) costs which we assume from now on imply that Y_t dynamically separates the two monopolies held by S_1 and S_2. A sketch of the situation is given in Fig. 2. As Y_t is a function of the two stochastic processes $N_1(t)$ and $N_2(t)$, it will be itself *a stochastic process.*

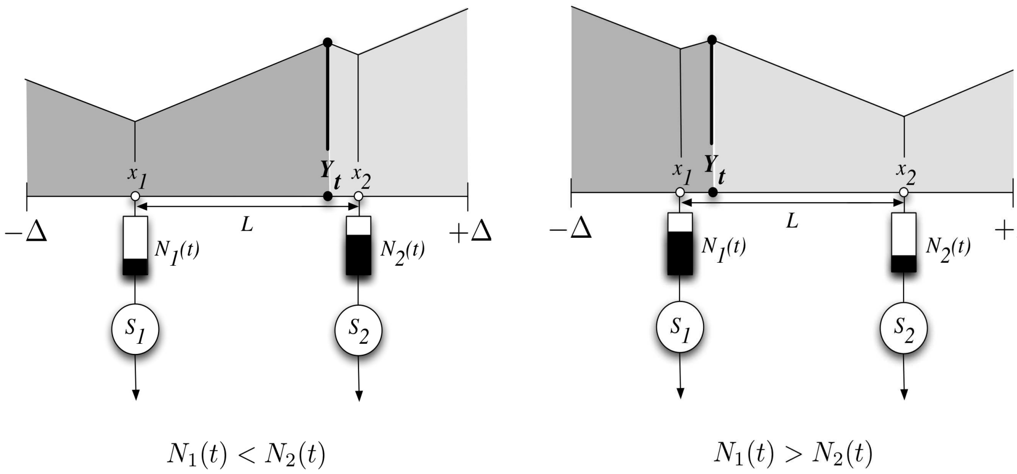

Figure 2: Cost structure in function of the customers' location. The total costs for a customer located at position x are the sum of the selling price p, the waiting time cost $c_w \mathbb{E}(W_i|N_i(t))$ and the brand departure cost $c_t|x - x_i|$. Any customer will choose the service provider which minimizes his/her total costs (i.e. it corresponds to maximize his/her utility function). As a consequence, all the customers standing on the left of Y_t will choose S_1, those on the right will choose S_2. The only difference between the two figures above is the current content of the queues. We clearly see that the values of these contents act upon the position of the boundary point Y_t, which separates the respective market shares held by S_1 and S_2.

Let $\lambda_i(t, Y_t)$ denotes the partial incoming rate of customers feeding S_i at time t and hence:

$$\lambda_1(t, Y_t) + \lambda_2(t, Y_t) = \Lambda, \qquad \forall t \in \mathbb{R}^+. \tag{3}$$

In view of the assumption b) (i.e. spatially uniform arrival on $\Omega = [-\Delta, +\Delta]$), the partial traffic flows feeding S_1 and S_2 result from the Bernoulli "thinning" of the incoming Poisson flow with global rate Λ. The branching probability is given by $P = \frac{\Delta - Y_t}{2\Delta}$ and it is well known that the thinning produces two independent Poisson processes with partial rates:

$$\lambda_1(t, Y_t) = \frac{\Delta + Y_t}{2\Delta}\Lambda \qquad \text{and} \qquad \lambda_2(t, Y_t) = \frac{\Delta - Y_t}{2\Delta}\Lambda. \tag{4}$$

Let $A_i(t)$, $D_i(t)$ and $N_i(t)$ respectively denote the numbers of arrivals, departures and the population in S_i at time t. From now on, we restrict ourselves to *heavy traffic* regimes characterized by $\rho = \frac{\Lambda}{2\mu} = 1 - \epsilon$, with ϵ small. Writing

$$N_i(t) = A_i(t) - D_i(t),$$

in heavy traffic the server S_i has very long busy period and hence the process $N_i(t)$ does almost never vanish, $i = 1, 2$. This implies that the departure and arrival processes are almost independent. In heavy traffic regimes, it is well established (see in particular [Mehdi 1991]) that both queue contents at time t are well approximated by diffusion processes of the form:

$$N_i(t) = \int_0^t [\lambda_i(s, Y_s) - \mu]\, ds + \int_0^t V_i(s, Y_s) dB_{i,s} \quad i = 1, 2, \tag{5}$$

where $B_{1,t}$ and $B_{2,t}$ are independent standard Brownian motions and the terms $V_i(t, Y_t)$ denote the state-dependent "volatilities" given by:

$$V_i(t, Y_t)^2 = \lambda_i(t, Y_t)^3 \sigma_{a,i}^2 + \mu^3 \sigma_{s,i}^2 \quad i = 1, 2, \tag{6}$$

with $\sigma_{a,i}^2$ (resp. $\sigma_{s,i}^2$) being the variance of the inter-arrival times (resp. the variance of the service times) for server S_i. Using Eqs.(3) to (6) and the fact that $B_{1,t}$ and $B_{2,t}$ are independent, we therefore can write:

$$N_2(t) - N_1(t) = -\frac{\Lambda}{\Delta} \int_0^t Y_s ds + \int_0^t V(s, Y_s) dB_s, \tag{7}$$

with B_t being a standard Brownian motion and $V^2(t, Y_t) = V_1(t, Y_t)^2 + V_2(t, Y_t)^2 = \Lambda + \mu^3 \left(\sigma_{s,1}^2 + \sigma_{s,2}^2\right)$ - remember that for Poisson processes, we have $\sigma_{a,i}^2 = \lambda_i(t, Y_t)^{-2}$.

When the utility functions are given by Eq.(1), the time-dependent boundary point will obey, $\forall t \in \mathbb{R}^+$:

$$Y_t = \begin{cases} \frac{c_w}{2\mu c_t} (N_2(t) - N_1(t)) & \text{if} \quad c_t L \geq \frac{c_w}{\mu} |(N_2(t) - N_1(t))| \\ -\Delta & \text{if} \quad c_t L < \frac{c_w}{\mu} (N_2(t) - N_1(t)) \\ +\Delta & \text{if} \quad c_t L < \frac{c_w}{\mu} (N_1(t) - N_2(t)) \end{cases} \tag{8}$$

Note that when $c_t L \geq \frac{c_w}{\mu} |(N_2(t) - N_1(t))|$, then $Y_t \in [x_1, x_2] \subset [-\Delta, +\Delta]$. Indeed in this case, the brand departure cost from one server to the other (i.e. $c_t L$) is greater than the difference between the waiting time costs of the two servers (i.e. $\frac{c_w}{\mu}|(N_2(t) - N_1(t))|$). Hence, a customer located near the server having the longest queue will choose this server anyway. When $c_t L < \frac{c_w}{\mu}|(N_2(t) - N_1(t))|$, then any customer in the whole interval $[-\Delta, +\Delta]$ will join the server having the shortest queue. Indeed, the difference between the content of the queues is such that the relative gain in waiting time cost is greater than the brand departure cost from one server to the other. A representation of the dynamics induced by Eq.(8) for a particular choice of the control parameters is found in Fig. 3.

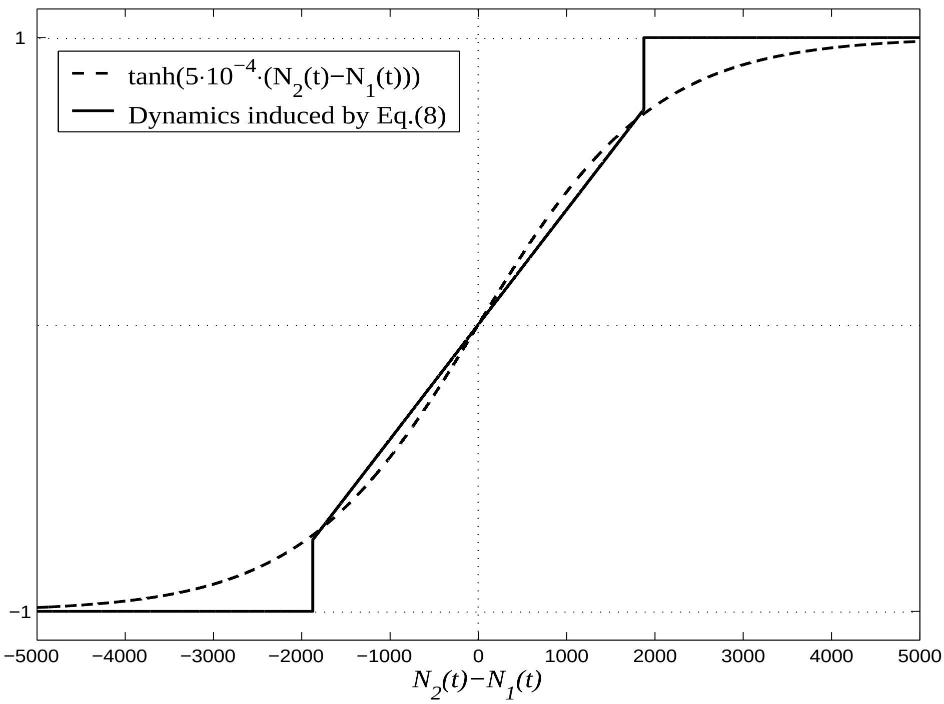

Figure 3: Particular representation of the boundary position dynamics when $\Delta = 1$. The solid line shows the dynamics induced by Eq.(8) when $c_t = 10$, $c_w = 8 \times 10^{-3}$, $\mu = 1$ and $L = \frac{3}{2}$. The dashed line shows the dynamics given by Eq.(20) when $\gamma = 5 \times 10^{-4}$.

To proceed further with analytical calculations and approximate the dynamics implied by Eq.(8), we introduce an odd (due to the symmetry of the problem), effective monotonously increasing one-to-one, $C^2(\mathbb{R})$ function:

$$f(\cdot) \; : \; \mathbb{R} \rightarrow [-1, +1] \tag{9}$$

fulfilling:

$$Y_t = \Delta f \left(\gamma \left(N_2(t) - N_1(t)\right)\right), \tag{10}$$

with:

$$\gamma := \frac{c_w}{\mu \, L \, c_t} \tag{11}$$

being a non-dimensional parameter quantifying the respective importance of the different costs. Note that in Eq.(11), the time unit is measured in average service time. A formal characterization of the functions which could be considered in the present model is given in the appendix.

As f is invertible, Eq.(10) can be written as:

$$f^{-1}\left(\frac{Y_t}{\Delta}\right) = \gamma \left(N_2(t) - N_1(t)\right). \tag{12}$$

Using Eq.(7), Eq.(12) becomes:

$$f^{-1}\left(\frac{Y_t}{\Delta}\right) = -\frac{\gamma\Lambda}{\Delta}\int_0^t Y_s ds + \gamma \int_0^t V(s, Y_s) dB_s. \tag{13}$$

Differentiating, we obtain:

$$\left(f^{-1}\right)'\left(\frac{Y_t}{\Delta}\right) dY_t = -\gamma\Lambda Y_t dt + \Delta\gamma V(t, Y_t) dB_t, \tag{14}$$

which can be written as:

$$dY_t = -\frac{\gamma\Lambda Y_t}{\left(f^{-1}\right)'\left(\frac{Y_t}{\Delta}\right)} dt + \frac{\Delta\gamma V(t, Y_t)}{\left(f^{-1}\right)'\left(\frac{Y_t}{\Delta}\right)} dB_t. \tag{15}$$

In our settings (remember that we deal with M/G/1 queues), $V(t, Y_t) = V = \sqrt{\Lambda + \mu^3\left(\sigma_{s,1}^2 + \sigma_{s,2}^2\right)}$ does not depend on Y_t nor on t. We can thus write Eq.(15) as:

$$dY_t = -\frac{\gamma\Lambda Y_t}{\left(f^{-1}\right)'\left(\frac{Y_t}{\Delta}\right)} dt + \frac{\Delta\gamma V}{\left(f^{-1}\right)'\left(\frac{Y_t}{\Delta}\right)} dB_t. \tag{16}$$

The stochastic differential equation (SDE) given by Eq.(16) describes the effective dynamics of the boundary position Y_t. The White Gaussian noise dB_t being merely the limit of finitely correlated processes, we assign to the underlying stochastic integral relative to Eq.(16) the Sratonovitch's interpretation. Hence, the transition probability density $P(y, t \mid y_0, t_0)$ describing the solution of the SDE (16) reads as:

$$\frac{\partial}{\partial t} P(y, t \mid y_0, t_0) = \mathcal{F} P(y, t \mid y_0, t_0), \tag{17}$$

with Fokker-Planck operator taking here the form, [Horsthemke et al. 1984]:

$$\mathcal{F}(\cdot) := \frac{\partial}{\partial y}\left[\frac{\gamma\Lambda y}{(f^{-1})'\left(\frac{y}{\Delta}\right)}\,(\cdot)\right]+\frac{1}{2}\frac{\partial}{\partial y}\left[g(y)\frac{\partial}{\partial y}g(y)\,(\cdot)\right], \quad g(y)=\frac{\Delta\gamma V}{(f^{-1})'\left(\frac{y}{\Delta}\right)}.$$

The stationary probability density function $P_s(y)$ solving Eq.(17), with vanishing left hand side, reads as:

$$P_s(y) = \mathcal{N}\left(f^{-1}\right)'\left(\frac{y}{\Delta}\right)\exp\left\{-\frac{2\Lambda}{\gamma\Delta^2 V^2}\int^y u\left(f^{-1}\right)'\left(\frac{u}{\Delta}\right)du\right\}, \quad (18)$$

for $y \in [-\Delta, +\Delta]$, with $\mathcal{N} < \infty$ a normalization constant.

Our assumptions of identical prices and identical dynamics of the servers imply an even parity of the stationary measure (i.e. $P_s(y) = P_s(-y)$). In particular, studying the curvature $\mathcal{R}(0)$ of $P_s(y)$ at $y = 0$ directly furnishes information regarding the modularity of $P_s(y)$. From Eq.(18), we directly obtain:

$$\text{sign}\left\{\mathcal{R}(0)\right\} = \text{sign}\left\{-\gamma V^2 f'''(0) - 2\Lambda\left(f^{-1}\right)'(0)\left(f'(0)\right)^3\right\}. \quad (19)$$

For given functions f, we observe that the sign of the curvature $\mathcal{R}(0)$ directly depends on the values of the (control) external parameters (here c_w, c_t, L, Λ and μ) of our class of models. This clearly shows the possibility to observe **noise-induced phase transitions** and an explicit illustration is worked out in section 3 to follow.

3 Explicit Illustration

Belonging to the previous class of models, the particular choice

$$Y_t = \Delta\tanh\left(\gamma\left(N_2(t) - N_1(t)\right)\right) \quad (20)$$

leads to very simple algebra. A particular representation of Eq.(20), put into comparison with the dynamics induced by Eq.(8), is found in Fig. 3.

For this particular case, the SDE (16), describing the effective boundary point dynamics, becomes:

$$dY_t = -\gamma\Lambda Y_t\left(1-\left(\frac{Y_t}{\Delta}\right)^2\right)dt + \Delta\gamma V\left(1-\left(\frac{Y_t}{\Delta}\right)^2\right)dB_t. \quad (21)$$

In view of Eq.(18), the corresponding stationary probability density function simply becomes:

$$P_s(y) = \mathcal{N}\left(1 - \left(\frac{y}{\Delta}\right)^2\right)^{\frac{\Lambda}{\gamma V^2}-1} \quad \text{for} \quad y \in [-\Delta, +\Delta], \tag{22}$$

where $\mathcal{N}$ is the normalization constant given here by:

$$\mathcal{N}^{-1} = \Delta \int_0^1 t^{-\frac{1}{2}} (1-t)^{\frac{\Lambda}{\gamma V^2}-1} dt = \Delta B\left(\frac{1}{2}, \frac{\Lambda}{\gamma V^2}\right),$$

where $B(x, y) := \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$ and $\Gamma(x)$ stands for the Gamma function. An illustration of the probability density function given by Eq.(22) for different values of γ and $\Delta = 1$ is found in Figure 4. Regarding

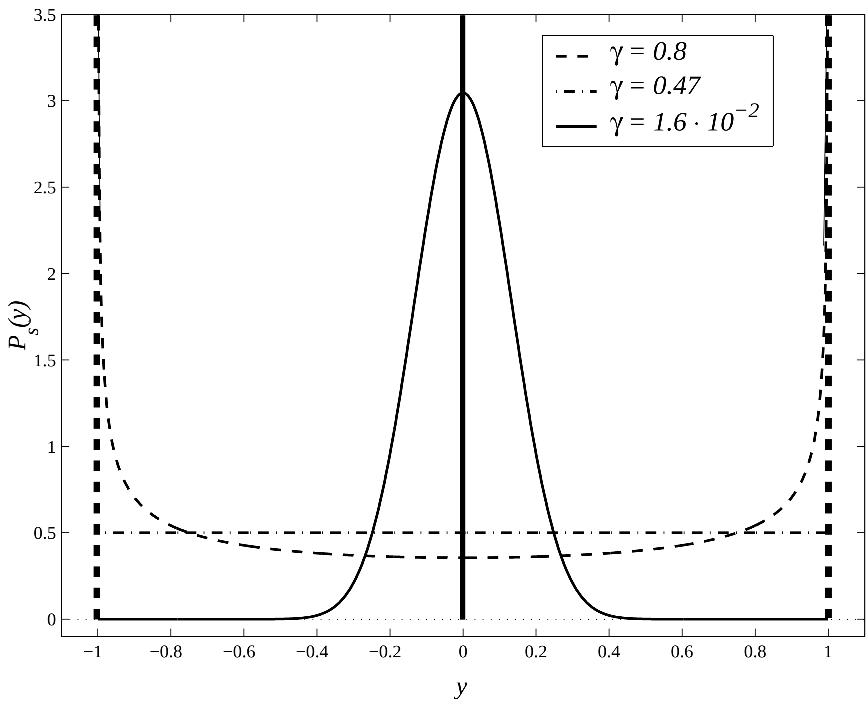

Figure 4: Stationary probability density function of the time-dependent boundary position Y_t when $\Delta = 1$, $\Lambda = 1.8$, $\mu = 1$ ($\rho = 0.9$) and the service time processes are Poisson. This density is drawn for three different values of $\gamma = [0.8; 0.47; 1.6 \cdot 10^{-2}]$. Furthermore, when $\gamma \to \infty$ (it corresponds to purely deadline type regimes), the density is sharply peaked at $y = -\Delta = -1$ and $y = +\Delta = +1$. In the other limit, $\gamma \to 0$ (corresponding to purely Hotelling-like regimes), the density is restricted to a single peak at $y = 0$. This graph clearly exhibits the noise induced phase transition arising in our dynamic model.

Eq.(19), the sign of the curvature $\mathcal{R}(0)$ of $P_s(y)$ at $y = 0$ is here given by:

$$\mathcal{R}(0)\begin{cases} >0 & \text{when} \quad \frac{\Lambda}{\gamma V^2}<1, \\ =0 & \text{when} \quad \frac{\Lambda}{\gamma V^2}=1, \\ <0 & \text{when} \quad \frac{\Lambda}{\gamma V^2}>1. \end{cases} \tag{23}$$

The information given by Eq.(23) (which is in perfect agreement with what we would expect with regard to the form of $P_s(y)$ given by Eq.(22)) perfectly describes the modularity of $P_s(y)$ and the underlying noise-induced phase transition.

3.1 Transient Behavior

For the choice given in Eq.(20), we can also study the rate of approach to the equilibrium. Indeed, by introducing the change of variables:

$$t \mapsto \tau = \gamma^2 V^2 t, \qquad X_t \mapsto Y_t = \Delta \tanh(X_t), \tag{24}$$

the dynamics given by Eq.(20) reduces to:

$$dX_\tau = -\frac{\Lambda}{\gamma V^2}\tanh(X_t) + dW_\tau := -2K\tanh(X_t) + dW_\tau \tag{25}$$

and the time-dependent solution $P(x, t \mid x_0, 0)$ of the associated Fokker-Planck is known for long (see for instance [Wong 1964]). As an illustration, let us mention that for the situations where the dimensionless parameter $K := \frac{\Lambda}{2\gamma V^2} \in \mathbb{N}$, the explicit form simplifies somewhat and is given by [Wong 1964]:

$$P(x,t \mid x_0, 0) = \frac{1}{1+z^2}\left[(1+z_0^2)(1+z^2)^{\frac{K}{2}}\frac{1}{2\sqrt{\pi\tau}}e^{-K^2\tau}e^{-\frac{(x-x_0)^2}{4\tau}}\right] +$$

$$\frac{1}{\pi(1+z^2)}\sum_{n=0}^{K-1}\frac{(K-n)}{n!\Gamma(2K+1-n)}e^{-n(2K-n)\tau}\theta_n(x_0)\theta_n(x)f_n(x,x_0,t), \tag{26}$$

with the definitions:

$$\sinh(z) := x, \qquad f_n(x,x_0,t) := \frac{1}{\sqrt{\pi}}\int_{\frac{(x-x_0)}{2\sqrt{t}}-(K-n)\sqrt{t}}^{\frac{(x-x_0)}{2\sqrt{t}}+(K-n)\sqrt{t}} e^{-z^2}\,dz$$

and the polynomials:

$$\theta_n(x) := (-1)^n 2^{K-n}\Gamma(K-n+\frac{1}{2})(1+x^2)^{K+\frac{1}{2}}\frac{d^n}{d^n}(1+x^2)^{n-K-\frac{1}{2}}.$$

In particular, the long time scale t_{relax} governing the approach to the stationary state given by Eq.(22) is determined by the spectral gap between 0 and the first non vanishing eigenvalue of the Fokker-Planck equation (17) (remember that the vanishing eigenvalue corresponds to the stationary probability measure given by Eq.(18)). It follows that:

$$1/t_{relax} = \begin{cases} (2K-1)\gamma^2 V^2 = \left(\frac{\Lambda}{\gamma V^2} - 1\right)\gamma^2 V^2 & \text{if} \quad K \geq 1, \\ \\ K^2\gamma^2 V^2 = \frac{\Lambda^2}{V^2} & \text{if} \quad K < 1. \end{cases} \tag{27}$$

From Eq.(27), we can draw the following remarks:

a) *Spectral characteristics of the Fokker-Planck equation.* In view of Eq.(27), there are two relaxation regimes governed by the spectral properties of the associated Fokker-Planck equation (17). As discussed in [Wong 1964], for $K \geq 1$ the spectrum exhibits both discrete and continuum parts whereas for $K < 1$ only the continuum part survives.

b) *Regime transitions.* Note that the transition from unimodal to bimodal densities given in Eq.(22) by $\left(\frac{\Lambda}{\gamma V^2} - 1\right) = 0$ coincides with the transition in the relaxation regimes given by Eq.(27)

c) *Rate of approach to the equilibrium.* When discrete eigenvalues exist, the asymptotic time relaxation towards the single mode stationary probability density (given by Eq.(22)) is faster compared to the relaxation rate associated with the purely continuum spectrum which drives the system to the bimodal density (given by Eq.(22)). This can be intuitively understood in limiting regimes. Indeed, note first that for the pure Hotelling case, the boundary position probability density is delta-peaked in the middle of the market interval, (remember that we did focus in this paper on fully symmetric configurations) and the relaxation time to reach this equilibrium is vanishingly small - this corresponds to the deterministic scheduling which commands to *"join the closest server"*. For dominating Hotelling's type regimes, the *externalities* (i.e. the waiting costs affecting incomers arriving behind a customer entering into service) have little influence on the equilibrium probability density which describes the boundary point - this produces a fast relaxation towards the statistical equilibrium, which will be close to the limiting delta-peaked density. In the contrary, when the deadline type regime strongly dominates, a new incomer strongly modifies the dynamical state of the system and hence strongly perturbs the underlying probability measure, thus implying a long relaxation time to the statistical equilibrium. Note that for $K = 0$ in Eq.(27), a situation realized when $c_t \to \infty$, the relaxation time diverges to infinity, meaning that no statistical equilibrium exists - this corresponds to the purely deterministic scheduling which commands to *"join the server exhibiting the shortest queue"*.

3.2 Simulation Experiments

We have simulated the dynamics of the boundary position Y_t in the particular case where Y_t fulfills Eq.(20). Simulations have been realized using the *Enterprise Dynamics* discrete events simulator. Each customer, upon arrival, determines on which side of the boundary point Y_t (dynamically given by Eq.(20), with regard to the current content of the queues) is his/her (uniformly distributed) position and he/she joins the queue hence chosen. We have computed an estimation of the stationary probability density function of the boundary position Y_t after 10^5 customers have passed through the system. The simulation experiments performed for different values of the control parameters (here γ, Λ and μ) confirm the presence of the noise-induced phase transition given by the analytical model. The particular result of such a simulation, put into comparison with the analytical curve given by Eq.(22), can be seen in Fig. 5.

4 Conclusion

Besides covering actual aspects of services, the addition of waiting costs to the original Hotelling's model confers dynamic and stochastic dimensions to a so far mostly static and deterministic elementary industrial organization problem. In the simplest configurations obtained for fixed and symmetric services, we already observe the central role played by the underlying random queue dynamics, which is here used to model the waiting processes. In particular for heavy traffic regimes, the Hotelling inner market boundary point obeys to a *time-dependent stochastic diffusion process* with multiplicative noise. Such multiplicative fluctuations, generated by state dependent "volatility" terms, are well known to give rise to *noise induced phase transitions*, a phenomena which cannot be derived by deterministic analysis alone.

Acknowledgments

This work is partially supported by the "FNSR" (Fonds National Suisse pour la Recherche) and by the Fundaçao para a Ciênca e a Tecnologica FCT, FEDER/POCTI-SFA-1-219, Portugal. The authors

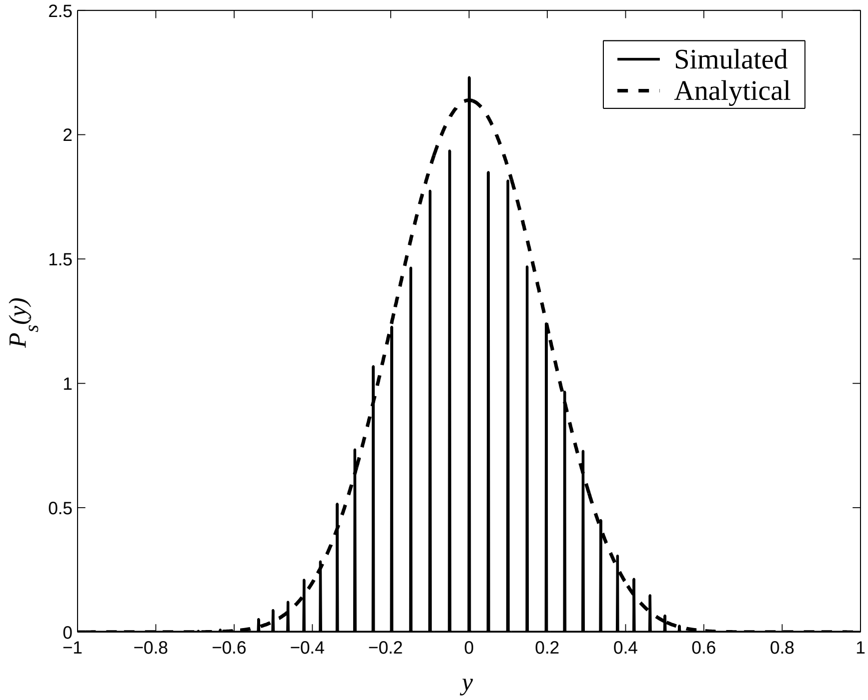

Figure 5: Simulated and theoretical stationary probability density function of the time-dependent boundary position $Y_t = \Delta \cdot \tanh(\gamma(N_2(t) - N_1(t)))$ when $\Delta = 1$, $\Lambda = 1.9$, $\mu = 1$ ($\rho = 0.95$), $\gamma = 5 \cdot 10^{-2}$ and the service time processes are Poisson.

would like to thank François Genoud for helpful discussions.

Appendix

It is possible to explicitly characterize the class of functions (which is very generally introduced in Eq.(9)) that can be considered in the present model in order to approximate the dynamics induced by the customers' decision policy. Indeed, we are interested in the functions for which there exists a particular value of γ (call it $\bar{\gamma}$) such that the stationary probability density function $P_s(y)$ given by Eq.(18) satisfies:

$$P_s(y) = \frac{1}{2\Delta} \quad \forall y \in [-\Delta, +\Delta]. \tag{28}$$

The condition given in Eq.(28) is necessary in order to be able to observe a phase transition between Hotelling-like and deadline type regimes. The unique class of functions fulfilling this condition is given in the following lemma.

Lemma 4.1. *The functions $f : \mathbb{R} \to [-1, +1]$ of the form:*

$$\left(f^{-1}\right)'\left(\frac{y}{\Delta}\right) = \frac{1}{\alpha\left(\frac{y}{\Delta}\right)^2 + \beta} \qquad \alpha < 0, \quad \beta \geq 0, \quad -\frac{\beta}{\alpha} \leq 1, \tag{29}$$

are the only ones for which there exists a particular value $\bar{\gamma} = -\frac{\Lambda}{\alpha V^2} > 0$ such that:

$$\begin{aligned} P_s(y) &= \mathcal{N}\left(f^{-1}\right)'\left(\frac{y}{\Delta}\right)\exp\left\{-\frac{2\Lambda}{\bar{\gamma}\Delta^2 V^2}\int^y u\left(f^{-1}\right)'\left(\frac{u}{\Delta}\right)du\right\} \\ &= \frac{1}{2\Delta} \quad \forall y \in [-\Delta, +\Delta]. \end{aligned} \tag{30}$$

For such functions f, $P_s(y)$ is either a unimodal or a bimodal distribution $\forall\ \gamma \neq \bar{\gamma}$.

Proof. First, note that Eq.(30) is satisfied iff $P_s'(y) = 0, \forall\, y \in [-\Delta, +\Delta]$. Fixing $\bar{\gamma}$ in Eq.(18) and differentiating, we obtain the following condition on f:

$$\left(f^{-1}\right)''\left(\frac{y}{\Delta}\right) - \frac{2\Lambda}{\bar{\gamma}\Delta V^2} y\left(\left(f^{-1}\right)'\left(\frac{y}{\Delta}\right)\right)^2 = 0 \quad \forall y \in [-\Delta, +\Delta]. \tag{31}$$

The unique set of solutions of this ordinary differential equation is given by:

$$\left(f^{-1}\right)'\left(\frac{y}{\Delta}\right) = \frac{1}{\alpha\left(\frac{y}{\Delta}\right)^2 + \beta} \qquad \beta \geq 0, \quad \forall y \in [-\Delta, +\Delta], \tag{32}$$

where $\alpha = -\frac{\Lambda}{\bar{\gamma}V^2} < 0$. Furthermore, in order that $f \in [-1, +1]$, we suppose that $-\frac{\beta}{\alpha} \leq 1$. Inserting Eq.(32) into Eq.(18) and using the fact that $\alpha = -\frac{\Lambda}{\bar{\gamma}V^2}$, we find that the stationary probability density function $P_s(y)$ is equal $\forall\ \gamma > 0$ to

$$\begin{aligned} P_s(y) &= \mathcal{N}\frac{1}{\alpha\left(\frac{y}{\Delta}\right)^2 + \beta}\exp\left\{-\frac{2\Lambda}{\gamma\Delta^2 V^2}\int^y u\frac{1}{\alpha\left(\frac{u}{\Delta}\right)^2 + \beta}du\right\} \\ &= \mathcal{N}\left(-\frac{\Lambda}{\bar{\gamma}V^2}\left(\frac{y}{\Delta}\right)^2 + \beta\right)^{\frac{\bar{\gamma}}{\gamma}-1} \qquad \forall y \in [-\Delta, +\Delta] \end{aligned} \tag{33}$$

for functions f satisfying Eq.(32). In regard to Eq.(33), $P_s(y)$ is hence either a unimodal or a bimodal distribution for $\gamma \neq \bar{\gamma}$. Note that for the function tanh ($\alpha = -1$ and $\beta = 1$), $\bar{\gamma} = \frac{\Lambda}{V^2}$ and Eq.(33) thus becomes:

$$P_s(y) = \mathcal{N}\left(1 - \left(\frac{y}{\Delta}\right)^2\right)^{\frac{\Lambda}{\gamma V^2}-1} \qquad \forall y \in [-\Delta, +\Delta],$$

which is in perfect agreement with Eq.(22). $\square$

Integrating Eq.(29) gives the following equivalent condition on f:

$$f^{-1}(y) = -\frac{1}{\sqrt{\alpha\beta}} \text{Arctanh}\left(\sqrt{-\frac{\alpha}{\beta}}y\right) + C, \quad C \in \mathbb{R}. \tag{34}$$

Then, determining the inverse of the functions satisfying Eq.(34), we find that the following monotonic increasing functions:

$$f(x) = \sqrt{-\frac{\beta}{\alpha}} \left(\frac{Ke^{\sqrt{-\alpha\beta}x} - e^{-\sqrt{-\alpha\beta}x}}{Ke^{\sqrt{-\alpha\beta}x} + e^{-\sqrt{-\alpha\beta}x}}\right), \quad K > 0, \tag{35}$$

satisfy Eq.(29) and hence, by Lemma 4.1, compose the class of functions we are interested in. Note that the function tanh belongs to this class ($\alpha = -1$, $\beta = 1$ and $K = 1$). Several functions belonging to the class characterized by Eq.(35) are drawn in Fig. 6. In addition to the

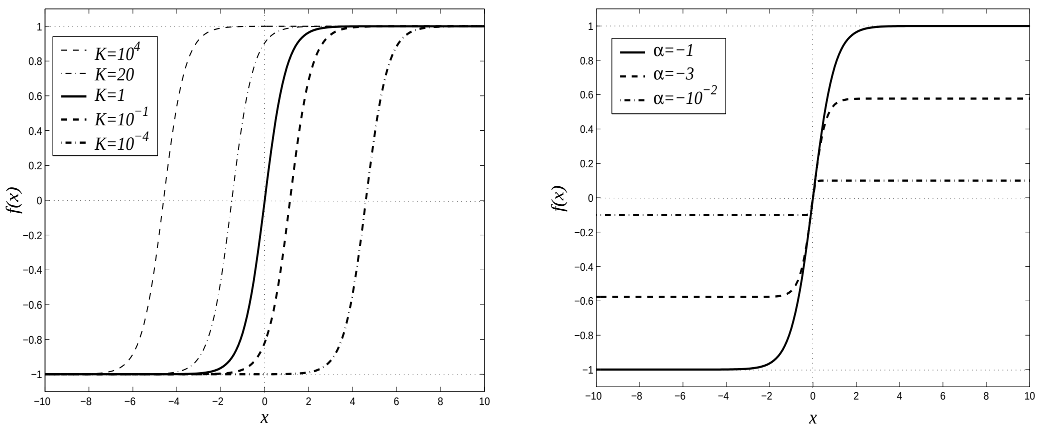

Figure 6: Several functions belonging to the class given by Eq.(35). *Left:* $\alpha = -1$, $\beta = 1$. *Right:* $\beta = 1$, $K = 1$.

large scope of dynamics covered by the choice of γ, the diversity of the functions resulting from the arbitrary choice of the tuple (α, β, K) in Eq.(35) allows us to cover a very wide range of dynamics induced by the customers' decision policy. Although we have restricted our study on purely symmetric cases, notice that it would also be possible to model non-symmetric cases by playing on the value of K (see Fig. 6, *Left*). Furthermore, the class of functions given by Eq.(35) would fit in order to model situations where the boundary point is restricted in a smaller interval $\mathcal{I} \subset \Omega$ (see Fig. 6, *Right*). This would correspond to the cases where a part of the customers are bound to a service provider, no matter the current state of the queue in front of it (for example when there is "competition" only in the interval $[x_1, x_2]$ between the two service providers, the customers in $[-\Delta, x_1]$ always choosing S_1 and the customers in $[x_2, +\Delta]$ always choosing S_2).

References

[1] G. P. CACHON and P. T. HARKER. *Service Competition, Outsourcing and Co-Production in a Queueing Game*, Preliminary Report, *http://knowledge.wharton.upenn.edu/papers/889.pdf*, March 1999.

[2] G. P. CACHON and P. T. HARKER. *Competition and Outsourcing with Scale Economies*, Management Science, Vol. 48, No. 10, pp. 1314-1333, October 2002.

[3] R. HASSIN and M. HAVIV. *To Queue or not to Queue*, Kluwer Academic Publishers, 2003.

[4] W. HORSTHEMKE and R. LEFEVER. *Noise-Induced Transitions*, Springer-Verlag, 1984.

[5] H. HOTELLING. *Stability in Competition*, The Economic Journal, Vol. 39, No. 153, pp. 41-57, March 1929.

[6] J. MEDHI. *Stochastic Models in Queueing Theory*, Academic Press, pp. 386-388, 1991.

[7] T. PUU. *Hotelling's "Ice cream dealers" with elastic demand*, The Annals of Regional Science, Vol. 36, pp. 1-17, 2002.

[8] E. WONG. *The Construction of a Class of Stationary Markoff Processes*, Sixteenth Symposium in Applied Mathematics - Stochastic Processes in Mathematical Physics and Engineering, ed. Richard Bellman, Providence, RI: American Mathematical Society, pp. 264-276, 1964.