A New Pricing Strategy for Restaurants: 
By-the-Weight with Entry Fee

Kiyoshi Yoneda  Alberto Matsuda  
{yoneda|alberto}@econ.fukuoka-u.ac.jp  
Fukuoka University Faculty of Economics

Abstract  A new pricing strategy is proposed which is a two-part markup pricing with a transaction cost almost as low as all-you-can-eat pricing. The strategy is an improvement of a by-the-weight pricing scheme popular in Brazilian business districts for serving lunch. The by-the-weight pricing has been introduced to Japan by Brazilian workers but unsuccessfully. The paper also proposes an explanation of why the same pricing scheme is successful in Brazil but not in Japan.

Keywords: restaurant, pricing, stress, behavior

1 Introduction

This paper proposes a new pricing strategy for restaurants which has the optimality of markup pricing together with a transaction cost almost as low as all-you-can-eat pricing. The strategy, called by-the-weight with entry pricing in this paper, is a slight modification of by-the-weight (or quilo or kilo) lunch pricing popular in Brazil.

In a restaurant the customer usually pays for what she ordered, as in the à la carte system, while in all-you-can-eat (smogasbord or
buffet) restaurants she can eat as much as she likes, perhaps in a
given amount of time. Yet other restaurants ask the customer to
pay a table charge (also cover charge or entrance fee) in addition to
the foods ordered. The last is a two-part tariff system

\[ p(q) = p_0 + p_1 q \]  

(1)

with

\[ p(q) : \text{total amount charged to the consumer}; \]
\[ p_0 : \text{entry fee}; \]
\[ p_1 : \text{price for unit consumption}; \]
\[ q : \text{quantity consumed}; \]

for which the consumer’s surplus can be maximized [4], while the
buffet and the à la carte without entrance fee are its special cases
with \( p_1 = 0 \) and \( p_0 = 0 \), respectively. The two-part tariff with
\( p_1 = 0 \) is sometimes called unlimited usage pricing while \( p_0 = 0 \) is
called linear pricing\(^1\) or proportional pricing.

The buffet pricing strategy has an advantage over the other two
when the transaction costs for the latter are high, as analyzed in [7].

The by-the-weight system mentioned above is an ubiquitous
pricing strategy for business lunch in Brazil: the restaurant charges
the customer by the weight of food she placed on her plate, in-
dependently of what she chose. A by-the-weight restaurant looks
almost exactly like an all-you-can-eat restaurant except that at the
end of serving tables a scale is waiting to weigh the foods on the
plate. The foods available vary among by-the-weight restaurants
because target customers range from street workers to executives.
This system is considered a form of price discrimination\(^2\) since the
same commodity (a unit weight of food) is produced with different
costs for different customers depending on the foods chosen.

\(^1\)The terminology “linear pricing” is misleading since \( p_0 + p_1 q \) with \( p_0 \neq 0 \)
is also said to be linear in mathematics.

\(^2\)Price discrimination [8, 10] usually means that the same commodity is sold
for different prices to different customers.
Although the origin of by-the-weight pricing is unclear, most Brazilians seem to accept the following story. Until the end of the 1970s most business lunch installations followed either the “set meal” (prato feito) or the all-you-can-eat model. First by-the-weight restaurants appeared in the 1980s simultaneously with the arrival of fast food chains. In a matter of a few years the system became popular throughout the country; today it is a nationwide standard for restaurants serving lunch.

By-the-weight restaurants are found not only in Brazil but also in neighboring South American countries such as Argentina, Uruguay, and Paraguay, though not as commonly as in Brazil. Since many Japanese descendants in South America come to work in Japan, by-the-weight restaurants are sometimes found in industrial areas with many Brazilian workers such as Gunma, Shizuoka, and Kanagawa prefectures. They are short-lived as a rule, soon falling back to charging different prices for different items. The research in this paper began out of curiosity on why the by-the-weight system, quite popular in Brazil, has been unsuccessful in Japan.

This paper clarifies the conditions under which the by-the-weight pricing system can be successful, explains why the system has been unsuccessful in Japan, and proposes an alternative which will potentially be successful.

The remainder of this paper is organized as follows. Section 2 briefly describes a general model of individual behavior based on a loss function, to be called the stress function, which generates a probability distribution of her behavior. Based on this general model, a simple specific model of consumer behavior is developed in section 3. In section 4 a typical consumer’s behavior is predicted using this model under the three pricing strategies, viz. markup,

---

3The initial idea was to exploit the price variation among foods, such as between meat and vegetables, which is much greater in Japan than in Brazil. The approach in this paper is simpler.

4A preliminary version of this model and a part of its analysis in the subsequent section appeared in [13].
buffet, and by-the-weight. After this analysis a new pricing strategy is proposed in section 5. Finally, section 6 concludes this paper with a possible direction for further research. Appendix 6 includes a numerical example.

2 A stress minimization model of individual behavior

Consider a system of weighted positive approximate equations

\[ x(\theta) \approx y \]  \hspace{1cm} (2)

\[ \theta := \begin{bmatrix} \vdots \\ \theta_j \\ \vdots \end{bmatrix} \hspace{1cm} x(\theta) := \begin{bmatrix} \vdots \\ x_i(\theta) \\ \vdots \end{bmatrix} \hspace{1cm} y := \begin{bmatrix} \vdots \\ y_i \end{bmatrix} \]  \hspace{1cm} (3)

\[ 0 < \theta_j \hspace{1cm} 0 < x_i(\theta) \hspace{1cm} 0 < y_i \]  \hspace{1cm} (4)

\[ \dim \theta \leq \dim x(\theta) = \dim y \]  \hspace{1cm} (5)

The first \( \dim \theta \) elements of \( x() \) are assumed to be the identity function, meaning that the corresponding equations are of the form

\[ x_i(\theta_j) = \theta_j \approx y_j \hspace{1cm} 1 \leq j \leq \dim \theta . \]  \hspace{1cm} (6)

In this paper \( x() \) will be linear; we will not try to state how general \( x() \) could be. The system may be thought of as representing requirements specifications: an individual adjusts her attitudes \( \theta \), which go through the mechanisms \( x() \) to produce actions \( x(\theta) \), trying to reach the targets \( y \) as closely as possible.

By letting

\[ z_i(\theta) := \frac{x_i(\theta)}{y_i} \]  \hspace{1cm} (7)
the equations $x(\theta) \approx y$ may be rewritten as $z(\theta) \approx 1$. We define the solution $\hat{\theta}$ of the system by

$$\hat{\theta} := \text{arg min}_\theta \varphi(\theta) \quad (8)$$

$$\varphi(\theta) := \sum w_is(z_i(\theta)) \quad (9)$$

where $s(z)$ is a smooth strictly convex function with $s(1) = 0$. The importance weights $w_i$, assumed given, represent a value system describing relative importance among the targets. The system of near-equations can easily be solved numerically since it reduces to an unconstrained minimization of a strictly convex function. This general model has been introduced in [12].

After Wichers [11] who uses a loss function instead of a utility function we call $f()$ the stress function. We prefer the loss function approach over the utility function approach for the same reason as Wichers, in [11] p.145:

"Why is it so easy to devise a justifiable specification of $s(z)$ [which is $\varphi(\theta)$ in the present paper] when it is so hard to do the same for [the utility function] $u(x)$? The answer does not lie in the difference between stress and utility. It lies, rather, in the difference between a minimand and a maximand. As scientists have demonstrated time and again, minimands tend to be easily specifiable. The reason is that they can often be given an elementary, physical interpretation, namely, that of a distance. Maximands lack an attractive feature of this sort. Nature abhors a maximand.

This points out the mathematical convenience that the value of stress for an attitude is immediately an indication of how far the result is from the aspired optimal since the level of stress is defined to be nonnegative. The same information is of course available from an utility function with a bliss point by subtracting its value from
the maximum utility which is usually normalized to one, viz. $1 - \text{utility}$, but this information is still a stretch of arm away.

While Wichers’s model requires a strict adherence to the budget constraints, the present model involves no hard constraints, at least in the short run\(^5\). This arises from the difference in the definition of stress: while Wichers’s stress function is a potential function to describe the evolution of an individual’s state with respect to time, our stress function is a statistical loss function \(^3\) which describes a probability distribution of the behavior of an individual in steady state (in stochastic systems usage) or equilibrium (in economics and physics usage), with no involvement of time. How our stress function is related to probability distributions is described in \([12]\).

From this probability distribution interpretation it follows that although our stress minimization approach may appear to be mathematically equivalent to utility maximization, there exists a fundamental difference between the two, as will be explained subsequently.

Suppose that a vector of attitude parameters $\hat{\theta}$ is optimal in the sense that the vector of actions $x(\hat{\theta})$ produced with $\hat{\theta}$ maximizes the utility function. Since $x()$ is a vector of deterministic functions, the prediction is such that the individual will always choose the same actions $x(\hat{\theta})$ under the same circumstances.

On the other hand, suppose a vector of attitude parameters $\hat{\theta}$ is optimal in the sense that it minimizes our loss function. Then it produces a representative vector of actions $x(\hat{\theta})$ which, in turn, determines a probability distribution of actions.

Although we are not going into the details of the probability distribution determined by the stress function because it will not be used in the present paper, we mention its relationship with the discrete choice theory \([1]\) in marketing and the matching law \([2]\) in experimental psychology. The discrete choice theory postulates

---

\(^5\)Another minor difference: while Wichers limits his stress to a quadratic function we leave its form open until necessity demands to be more concrete.
stochastic decision rules in which individuals do not necessarily choose the alternative that yields the highest utility but chooses any of the possible alternatives with a positive probability. This is in accordance with the matching law in psychology which models the experimentally established fact that an individual distributes choices over the available alternatives. Our loss function is consistent with the probability distribution the matching law proposes, called the generalized or modern matching law [6, 5].

3 Consumer response to pricing schemes

In this section we build a simple stress minimization model of consumer behavior in restaurants which takes into account only the quantity of food to consume and its price.

First we change the notation in the previous section for the sake of clarity in our specific problem:

\[ q := \theta_1 \] quantity actually consumed;

\[ \hat{q} := y_1 \] most preferred quantity to consume;

\[ p(q) := x_2(\theta) \] outlay for quantity \( q \);

\[ b(\hat{q}) := y_2 \] best outlay for \( \hat{q} \) available in the market;

\[ w_q := w_1 \] weight assigned to the quantity consumed;

\[ w_p := w_2 \] weight assigned to the total outlay.

Next we modify

\[ p(\hat{q}) \approx b(\hat{q}) \] (10)

to

\[ p(q) \approx b(q) \] , (11)
meaning that the consumer wants to pay the best price\(^6\) not only for the quantity \(\hat{q}\) but for any quantity \(q\) she actually consumes. This change renders the original bidimensional model with two variables, the quantity to eat and the amount to pay, into a monodimensional model with the sole variable \(q\). Then the requirements specification may be written as

\[
\begin{bmatrix} q \\ p(q) \end{bmatrix} \approx \begin{bmatrix} \hat{q} \\ b(q) \end{bmatrix}
\]

or

\[
\begin{bmatrix} q/\hat{q} \\ p(q)/b(q) \end{bmatrix} \approx \begin{bmatrix} 1 \\ 1 \end{bmatrix}
\]

with the stress function

\[
\varphi(q) := w_qs\left(\frac{q}{\hat{q}}\right) + wp\delta\left(\frac{p(q)}{b(q)}\right). \tag{14}
\]

Its derivative is

\[
\frac{d\varphi}{dq} = \frac{w_q}{\hat{q}} \frac{d}{dq} \left(\frac{q}{\hat{q}}\right) + \frac{w_p D(q)}{b(q)^2} \frac{d}{dq} \left(\frac{p(q)}{b(q)}\right) \tag{15}
\]

with

\[
D(q) := \left| \begin{array}{cc} b(q) & p(q) \\ \frac{db(q)}{dq} & \frac{dp(q)}{dq} \end{array} \right|. \tag{16}
\]

When the desired amount \(q = \hat{q}\) is consumed,

\[
\frac{d\varphi}{dq} (\hat{q}) = \frac{w_p D(\hat{q})}{b(\hat{q})^2} \frac{d}{dq} \left(\frac{\hat{q}}{b(\hat{q})}\right). \tag{17}
\]

Since \(0 < b(q) < p(q)\) and \(0 < w_p\), the sign of \(d\varphi(\hat{q})/dq\) depends only on that of \(D(\hat{q})\).

\(^6\)Lowest price available in the market.
4 Behavior prediction

4.1 How much we eat

Assume that our consumer has decided in which restaurant to eat so that the only remaining problem for her is the quantity to eat, \( q \). If the best market price is \( b(q) \) and what the consumer pays are both linear functions of \( q \)

\[
b(q) = b_0 + b_1 q
\]  \hspace{1cm} (18)

and

\[
p(q) = p_0 + p_1 q ,
\]  \hspace{1cm} (19)

then

\[
D := D(q) := b_0 p_1 - b_1 p_0
\]  \hspace{1cm} (20)

for all \( q \).

In the absence of a fee that depends on quantity consumed,

\[
p_1 = 0 \quad \Rightarrow \quad D < 0 \quad \Leftrightarrow \quad \frac{d\varphi(\hat{q})}{dq} < 0 ,
\]  \hspace{1cm} (21)

meaning that our consumer can reduce her stress \( \varphi(q) \) by eating more than her ideal amount \( \hat{q} \). This explains why we tend to overeat under the buffet pricing.

Likewise in the absence of an entry fee,

\[
p_0 = 0 \quad \Rightarrow \quad 0 < D \quad \Leftrightarrow \quad 0 < \frac{d\varphi(\hat{q})}{dq} ,
\]  \hspace{1cm} (22)

meaning that our consumer can reduce her stress \( \varphi(q) \) by eating less than her ideal amount \( \hat{q} \). This explains why we tend to stop short of enough in \( à la carte \) restaurants without entrance fee.
The pricing which protects the consumer from both under- and overeating is of course
\[ \frac{p_0}{p_1} = \frac{b_0}{b_1}, \quad (23) \]

viz. a two-part tariff.

Since
\[ \frac{p_0}{p_1} = \frac{b_0}{b_1} \Leftrightarrow p_0 + p_1 q = (1 + m)(b_0 + b_1 q), \quad 0 \leq m \quad (24) \]

and since the best price the suppliers can jointly offer under perfect competition equals the cost, the regular markup pricing should best protect the consumer from under- and overconsumption. From the restaurant’s point of view this pricing also protects the supplier by preventing the consumer from taking profit-reducing behavior, viz. under- or overeating.

We have established a way to predict the amount an individual actually eats in a restaurant, given her ideal amount to eat \( \hat{q} \), her importance weights \( w \), and the pricing strategy \( p(q) \) the restaurant adopts.

4.2 Where we eat

Now suppose that all three pricing strategies, markup, buffet, and by-the-weight offered the same total outlay for the ideal quantity \( \hat{q} \) for our customer to eat, as in Figure 1. Then from the discussion in the previous section it follows that for buffet and by-the-weight pricing the individual can always reduce her stress by eating more or less than the ideal amount \( \hat{q} \), respectively, while she cannot improve her stress for markup pricing, as illustrated in Figure 2.

Note that since a stress function determines a probability distribution of behavior, we can calculate the relative frequency an individual chooses among the three strategies. This is of course subject to the condition that all necessary data are available, the hardest to acquire being the importance weights \( w_q \) and \( w_p \).
Figure 1: Buffet, by-the-weight, and markup pricing schemes
Figure 2: Stress under buffet, by-the-weight, and markup pricing schemes
5 A new pricing strategy

The best pricing strategy which maximizes consumer surplus is the two-part tariff by markup pricing under which the marginal revenue equals the marginal cost [4]. In addition to this, we have seen that:

- The same markup pricing is the best for our restaurant customer in the sense that it induces neither under- or overeating.
- If the three pricing strategies — markup, by-the-weight, and all-you-can-eat — were presented simultaneously to our customer for the same fee at her optimal quantity to eat \(\hat{q}\), she has an incentive to choose the by-the-weight or the all-you-can-eat option to reduce her stress by under- or overeating, respectively.

Now we have a hypothesis why by-the-weight restaurants are unsuccessful in Japan: it cannot offer a price close to the markup scheme at the quantity most people want to eat, because the fixed cost is too high for the by-the-weight strategy to approximate the markup pricing well. See Figure 3.

Let the business total profit \(\pi\) be

\[
\pi = p(q) - c(q)
\]  
(25)

where \(c(q)\) is the cost function.

If

\[
c(q) = c_0 + c_1 q ,
\]  
(26)

where \(c_0\) is the fixed cost and \(c_1\) the cost of food, then the break-even point is

\[
q_e := \frac{-p_0 - c_0}{p_1 - c_1} .
\]  
(27)
Figure 3: Break-even-point $q_e$ in by-the-weight system
For by-the-weight restaurants where $p_0 = 0$, if the fixed cost $c_0$ is small the minimum amount a customer must consume in order to incur no loss to the restaurant is also small, so that the influence of under-consumption in the profitability of by-the-weight restaurants is small. Our hypothesis has been that this is what happens in Brazil. In contrast, when $c_0$ is significant, the said quantity is high, so that under-consumption is a considerable risk for by-the-weight restaurants. Our hypothesis has been that this is what happens in Japan.

A way to guarantee a supplier’s profit under any quantity the customer chooses is to add an entrance fee $c_0 < p_0$ to the by-the-weight pricing. The strategy (1) should not increase transaction cost significantly, and (2) sets the slope for $p(q)$ less steep than that for the by-the-weight pricing. This is our proposal. The proposal, by-the-weight with entry, is a markup pricing strategy such that the transaction cost is close to those in all-you-can-eat and by-the-weight schemes.

6 Conclusion

The simple by-the-weight pricing, popular in Brazilian business lunch, has been unsuccessful in Japan. An explanation of this phenomenon proposed in this paper has been that while in Brazil the by-the-weight pricing approximates the markup pricing well, fixed cost is too high in Japan for the same to happen. Based on this observation a new two-part pricing strategy, called the by-the-weight with entry, has been proposed for restaurants, which features low transaction cost, almost as low as with the by-the-weight and the all-you-can-eat pricing.

With the stress model of individual behavior it is possible to predict not only the quantity she eats under a given pricing scheme but also the frequency of use of restaurants, provided that her importance weights for quantity and its fee are available. Since impor-
tance weights are difficult to obtain in practice, a way to calculate them from easily available data is desirable. This would make the theory useful for practical marketing research.

Acknowledgments

We are grateful to Professor Babu Nahata, one of the authors of [7], who gave us lectures on the basics of pricing theory and reviewed an earlier version of this paper. We used R [9] for numerical experiments and to generate the graphs.

References


Appendix

Numerical example

This example illustrates how a customer behaves in response to the type of pricing strategy adopted. Our consumer makes her decision in two stages: first, whether or not to have lunch in a particular restaurant and second, how much to eat there. We assume that the consumer has already made the first decision, that she will eat in our restaurant. The consumer faces the second decision on the
quantity \( q \) to eat. She has a physically optimal amount to eat of 400 grams \([g]\). If the consumer eats less than 400 grams she remains unsatisfied; if she eats more, she feels uncomfortable for eating too much. The markup is set to \( m = 0.3 \) \( (p(400) = 780) \). We suppose that the three pricing strategies will give equivalent outlays at the ideal quantity \( \hat{q} \).

Let \( \hat{q} := 400 \ [g] \) with importance weights \( w_q := 2 \) for the quantity consumed and \( w_p := 3 \) for the total outlay, with

\[
\begin{align*}
b(q) &:= 200 + q \ \text{[yen]} \quad \text{cost or best market price;} \\
p_2(q) &:= 260 + 1.3q \ \text{[yen]} \quad \text{markup tariff;} \\
p_b(q) &:= 780 \ \text{[yen]} \quad \text{buffet tariff;} \\
p_k(q) &:= 1.95q \ \text{[yen]} \quad \text{by-the-weight tariff.}
\end{align*}
\]

Setting \( s(z) := (\log z)^2 \) the behavior is equivalent to the maximization of a utility function with a bliss point proportional to the lognormal density function. Then

\[
\begin{align*}
\varphi_2(q) &= 2 \left( \log \frac{q}{400} \right)^2 + 3 \left( \log \frac{260 + 1.3q}{200 + q} \right)^2 & q_{\varphi_{\min}} = 400 \\
\varphi_b(q) &= 2 \left( \log \frac{q}{400} \right)^2 + 3 \left( \log \frac{780}{200 + q} \right)^2 & q_{\varphi_{\min}} = 470 \\
\varphi_k(q) &= 2 \left( \log \frac{q}{400} \right)^2 + 3 \left( \log \frac{1.95q}{200 + q} \right)^2 & q_{\varphi_{\min}} = 355
\end{align*}
\]
Figure 4: $p(q)$ for buffet, markup, and by-the-weight schemes, all equivalent for $\hat{q} = 400$, $p(\hat{q}) = 780$
Figure 5: Stress $\varphi(\hat{q})$ under buffet (solid), markup (dashed) and by-the-weight (dotdash) schemes for $\hat{q} = 400$ and $s(z) := (\log z)^2$