Les Cahiers du GERAD ISSN: 0711–2440

Pricing decisions in fast fashion retailing using discrete choice dynamic programming model

E. Ergin
M. Gumus

G–2017–06

January 2017

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Pricing decisions in fast fashion retailing using discrete choice dynamic programming model

Elcin Ergin [a]

Mehmet Gumus [b]

[a] *Desautels Faculty Of Management, McGill University, Montréal (Québec) Canada*

[b] *GERAD & Desautels Faculty Of Management, McGill University, Montréal (Québec) Canada*

elcin.ergin@mail.mcgill.ca
mehmet.gumus@mcgill.ca

January 2017

Les Cahiers du GERAD

G–2017–06

Abstract: In this study, the problem environment consists of two fast fashion retailing firms where one can obtain the other's selling data from an outside agent. The aim of the study is to assess the value of such a third party and the each information piece obtained from it. This paper proposes a labelling algorithm based on a dynamic discrete choice modelling approach and uses it to study the effect of different attributes on estimation performance. Our extensive numerical study shows that the dynamic discrete choice-based procedure leads to highly accurate estimation for the hidden demand factors. Secondly, we show that if the seller is a strategic decision maker, it actually enhances the degree of accuracy as it becomes easier for the discrete-choice model to reverse-engineer the true values of the hidden demand factors. Finally, as the depth of discount increases, the degree of accuracy decreases. This comes from the fact that the deeper discounts disincentivizes the seller to offer discounts, which in turn reduces the chances of observing the demand at different price points.

Keywords: Retailing, fast fashion, pricing, dynamic discrete choice models

1 Introduction

In fast fashion retailing, rapid and effective translation of the changes in fashion trends into the selling strategies is a necessity for the success. The leading companies are the ones who manage to foresee the upcoming trends and adjust their assortment, inventory and pricing decisions accordingly. One of the world's leading companies, Zara, is well known by their successful business model where they use sophisticated operations research methods which enables them to efficiently lead their operations (Caro et al., 2010). In the market, while the big actors benefit from their advanced tools and flexible supply chain systems, smaller and local actors cannot easily set up such systems since it might not be feasible for them. However; as they still need to get the knowledge on current trends in fast fashion sector, they look for alternative ways to obtain this information. In some markets, there are third party companies which collect competitors' selling data and provide to the firm that they work for. For the smaller firms, such data could be useful to depict the information that competitor firms already have. They can examine competitors' pricing decision process and find out the driving forces for them to make these decisions.

In this study, we aim to assess the value of the competitor firms' data that third party companies provide to a fast fashion retail company.

Research question 1: Can the hidden information known only by the competitors be identified?

Research question 2: How the content of this data affect the amount of information that the firm can depict?

Research question 3: Under which conditions the identification error increases or decreases?

In order to answer these research questions, we use dynamic discrete choice structural model because of two reasons. First of all, this model enables us to address the dynamic nature of pricing actions. We can examine the effect of each pricing decision on the upcoming decisions will be made throughout the the given time period. Secondly, the estimation is done under the principle of revealed preference in this model (Aguirregabiria and Mira, 2010). Even if a specific action is not seen in the data, knowing that another action is chosen over the other one, the model can depict some information related to the unobserved action.

The dynamic discrete choice framework is widely applied on various problems in the literature such as fertility modelling (Wolpin, 1984), job matching (Miller, 1984), patent renewal (Pakes, 1984), bus engine replacement (Rust, 1987), career decision modelling (Keane and Wolpin, 1997), brand choice (Erdem et al., 2003), product launch (Hitsch, 2006) etc. The main advantage of this framework is the tractability provided for the choice models. However, the tradeoff is the computational burden comes from two sources which are curse of dimensionality and complexity of likelihood function. In order to handle this computational issue, papers such as (Hotz and Miller, 1993), (Keane and Wolpin, 1994), (Rust, 1997), (Aguirregabiria and Mira, 2002) and (Ackerberg, 2009) proposed approximate solution methods for approximate dynamic programming. Recently (Imai et al., 2009) utilized Bayesian approach to significantly decrease the computational burden for infinite horizon dynamic discrete choice problems. (Ishihara and Ching, 2016) modifies their method to apply a bayesian approach for finite horizon problem. They propose a modified backward reduction approach to reduce the number of value function computation.

The remainder of the paper is structured as follows. In Section 2 we formulate the problem and assumptions imposed in this paper. In Section 3, we explain the use of dynamic discrete choice models as an estimation algorithm. Section 4 presents the results of the numerical study. Finally, we conclude with future research directions in Section 5.

2 Problem formulation

We consider a market where two fast fashion retail companies are operating and selling a specific product. Firms sell the product throughout a given season. Firm 1 has more information on the demand structure of the product and starts to sell it before Firm 2. As seen in Figure 1 firms start the season by determining inventory levels for the product according to demand information they have. It is assumed that no inventory replacement is made during the season. At each period, firms choose their selling prices as regular or discounted price. It is assumed that there are two predetermined price levels. Firms want to maximize their expected total profit. Firms can estimate their expected profit either myopically or strategically. We assume Firm 1 starts selling the product earlier than Firm 2. Firm 2 obtains various information related to Firm 1's selling process from the outside agent since Firm 2 wants to estimate the part of the demand known by Firm 1. Hence, we focus on Firm 1's selling process throughout this study.

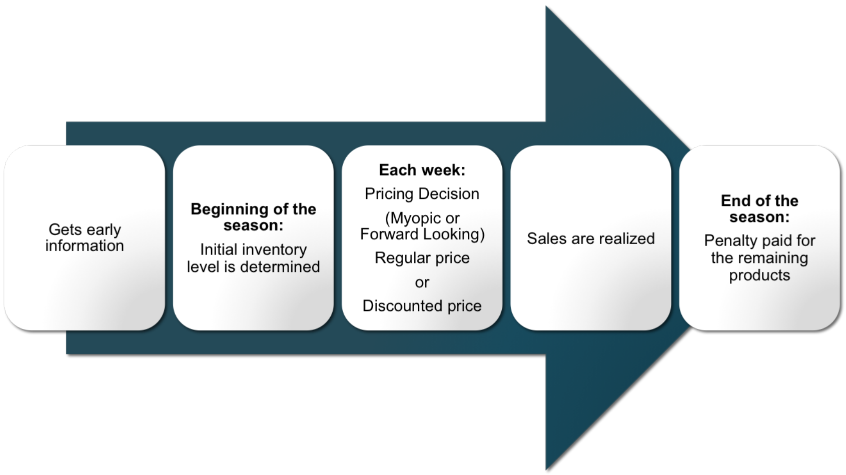

Figure 1: Process of a firm in the market

In this study, due to the two-level pricing assumption, we also define two levels for the demand structure corresponding to regular and discounted price as in Equation 1 where each level consists of three parts which is shown in Equation 2 and 3. First part is the base demand, represented by μ, which is exactly known by both firms. Both firms are able to use their own historical data and experience to extract this information. Our focus in this study is the second parameter γ, which represents the finer demand information only known by Firm 1. We also assumed that there are two levels for this part representing whether the product is a trendy or an ordinary product. Firm 2 only has a priori information. In this case, the values of the parameters at corresponding type of the product are known by Firm 2 but the true type is not known. Finally, demand also includes a random shock part, represented by ϵ, whose true level is not known by both firms. Random shock is assumed that normally distributed with zero mean and σ standard deviation.

$$D_t = \begin{cases} D_R & p_t = p_R \\ D_D & p_t = p_D \end{cases} \tag{1}$$

$$D_R = \overbrace{\mu_R}^{\text{Known by everyone}} + \overbrace{\gamma_R^m}^{\text{Only known by the seller}} + \overbrace{\epsilon}^{\text{Random Shock}} \tag{2}$$

$$D_D = \overbrace{\mu_D}^{\text{Known by everyone}} + \overbrace{\gamma_D^m}^{\text{Only known by the seller}} + \overbrace{\epsilon}^{\text{Random Shock}} \tag{3}$$

where $\epsilon \sim N(0, \sigma^2)$, $m = trendy\ or\ m = ordinary$. We also assume that

$$\gamma_f^{\text{trendy}} \geq \gamma_f^{\text{ordinary}} \tag{4}$$

for all $f \in \{R, D\}$.

State variables represent the additional information that Firm 1 uses while making their pricing decisions at each period. Because of tractability and computational issues, two state variables are kept which are selling price level at previous period and number of units on hand at the beginning of the current period. Specifically, Firm 2 obtains these state variables from the outside agent and try to estimate the type of the product.

States:

s_{1t}**:** Chosen price at t-1

s_{2t}**:** Number of units in the inventory at the beginning of period t

Action:

p_t**:** Price of the product in period t

$p_t \epsilon \{p_R, p_D\}$

Firm 1's price optimization

Firm 1 makes pricing decision at the beginning of each period in order to maximize the expected profit values. We represent this as a dynamic model where this decision process is repeated at every period during the selling season of the product. The firm calculates expected profit which will be generated both under the action of regular price and discounted price as in Equation 5. Then, the firm chooses the price level for the corresponding period. Next period, previous period's price action becomes a new state variable and the firm does these calculations again.

$$V_t(s_{1t}, s_{2t}) = max(E_{D_R}[V_t^R(s_{1t}, s_{2t})], E_{D_D}[V_t^D(s_{1t}, s_{2t})]) \tag{5}$$

For the firm, value function calculations can change according to how much future periods' returns are considered. The firm could be interested in only the current period's expected profit or they also look at the upcoming periods. This study examines two extreme cases where the firm is myopic or completely forward looking. By "completely forward looking", it is meant that the firm's value function includes the expected value functions for all the periods in the remaining part of the season. In practice, firms might not be distinctively myopic or forward looking. Some firms might be more forward looking than others or vice versa. Still, examining these two distinct cases would give a broad understanding of the effect of the level of firm's forward looking behaviour.

Case I: *Firm 1 is myopic*

$$V_t^R(s_{1t}, s_{2t}|D_R) = p_R * [min(s_{2t}, D_R)]^+ \tag{6}$$

$$V_t^D(s_{1t}, s_{2t}|D_D) = p_D * [min(s_{2t}, D_D)]^+ \tag{7}$$

If the firm is myopic, the value functions are calculated as in Equations 6 and 7 for regular and discounted price options, respectively. The firm can sell number of products demanded at the corresponding price level as long as the inventory on hand allows for it. The firm only calculates the revenue they obtain from these sales.

Case II: *Firm 1 is dynamic*

$$V_t^R(s_{1t}, s_{2t}|D_R) = p_R * [min(s_{2t}, D_R)]^+ + V_{t+1}(p_R, [s_{2t} - D_R]^+) \tag{8}$$

$$V_t^D(s_{1t}, s_{2t}|D_D) = p_D * [min(s_{2t}, D_D)]^+ + V_{t+1}(p_D, [s_{2t} - D_D]^+) \tag{9}$$

If Firm 1is forward looking, the value functions are calculated recursively as it is shown in Equations 8 and 9.

3 Estimation algorithm

The estimation algorithm constructed in this study is a labelling algorithm which aims to reveal the product type that can be either trendy or ordinary. Basically, the algorithm constructs likelihood values for two possible product types. As explained in Section 1, our algorithm is based on dynamic discrete choice approach. At each period, the algorithm looks at the actual chosen price by Firm 1 and compare expected value functions for regular and discounted price both under the product is trendy and ordinary. Then according to consistency of the actual chosen price by Firm 1 and this comparison, likelihood values are calculated for each type. In Figure 2, estimation procedure is shown as process chart.

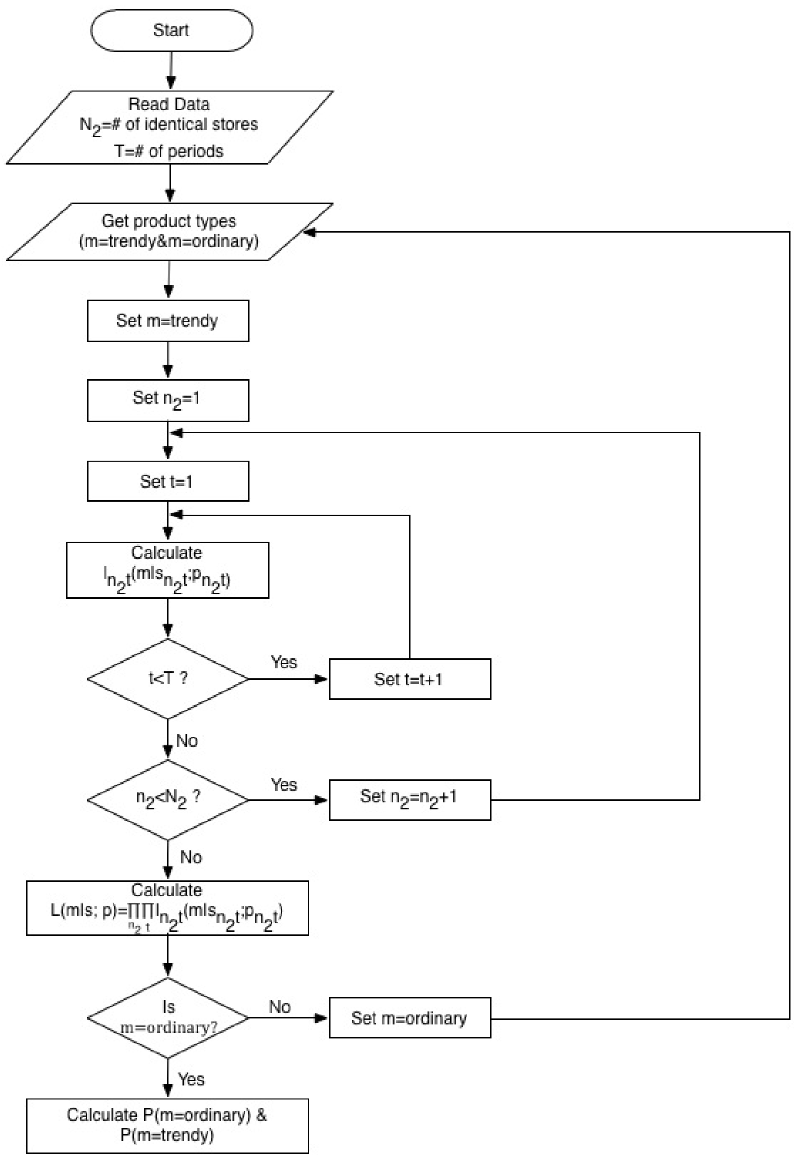

Figure 2: Estimation procedure

The probability of observing choice j (regular or discounted) given the observed state $s_t = \{s_{1t}, s_{2t}\}$ and assuming type of the product is m is presented in Equation 10. This is the probability of expected profit generated by choosing price j being higher than the expected profit generated by choosing any other possible price level.

$$P_j(s_t; m) = P(E_{D_j}[V_t^j(s_t; m)] \geq E_{D_k}[V_t^k(s_t; m), \forall k)] \tag{10}$$

Since we discretize the finer demand part for the product by defining two types and actions in two levels, the likelihood increments has also a discretized structure as shown in Equation 11. Here $l_t(m|s_t; p_t)$ represents the likelihood increments for type m under given states and actual chosen price p_t by Firm 1 at period t.

$$l_t(m = j|s_t; p_t) = \begin{cases} 1 & E_{D_j}[V_t^j(s_t; m)] > E_{D_k}[V_t^k(s_t; m)] \& p_t = j \\ \frac{1}{2} & E_{D_j}[V_t^j(s_t; m)] = E_{D_k}[V_t^k(s_t; m)] \\ 0 & otherwise \end{cases} \tag{11}$$

The likelihood of the product type m is equal to j is calculated by summing all the likelihood increments at each period as in Equation 12. Finally, using these likelihood probabilities the probability of the product type is j is found by normalization as shown in Equation 13.

$$L(m = j|s; p) = \sum_t l(m = j|s_t; p_t) \tag{12}$$

$$P(m = j) = \frac{L(m = j|s; p)}{\sum_i L(m = i|s; p)} \tag{13}$$

Table 1 represents how the algorithm works on a simple 4-week example where it is assumed that the true product type is trendy. The available data includes states and pricing decisions made by Firm 1. At each period, a consistency check is applied using expected value functions. For the first week, we assume that according to the consistency check, agent would earn more profit by offering regular price both for a trendy and ordinary product. These results are consistent with Firm 1's action since regular price is seen in the first week. Since the consistency check gives the same results with the chosen action for the both types, both likelihoods are incremented by one. In the second week, under the assumption that product type is ordinary, consistency check fails since according to calculations choosing discounted price would yield higher profit which contradicts with the Firm 1's action. Hence the likelihood increments for the product type is ordinary does not get any value. In the case of equality, considering that there is not a strong superiority of an action over another, the likelihood increment takes a value of 0.5. The remaining of the likelihood increments are calculated in the same fashion and these increments constitutes the likelihood values. Finally with a simple normalization procedure probabilities for each product type is calculated. The algorithm labels the product type according to these probabilities.

Table 1: Estimation algorithm demonstration example

Week	State	Price	Consistency Check	$\mathbf{l_t(m = tr)}$	$\mathbf{l_t(m = or)}$
1	S_1	R	$V_1^R(S_1; m = tr)] > V_1^D(S_1; m = tr)$ $V_1^R(S_1; m = or)] > V_1^D(S_1; m = or)$	1	1
2	S_2	R	$V_2^R(S_1; m = tr)] = V_2^D(S_1; m = tr)$ $V_2^R(S_1; m = or)] < V_2^D(S_1; m = or)$	1/2	0
3	S_3	D	$V_3^R(S_1; m = tr)] = V_3^D(S_1; m = tr)$ $V_3^R(S_1; m = or)] > V_3^D(S_1; m = or)$	1/2	0
4	S_4	D	$V_4^R(S_1; m = tr)] < V_4^D(S_1; m = tr)$ $V_4^R(S_1; m = or)] < V_4^D(S_1; m = or)$	1	1
Sum				3	2
Prob				3/5	2/5

4 Numerical results

4.1 Design of experiment

In order to test the estimation procedure, a simulation study is conducted. Selling data of Firm 1 is simulated under a set of parameters represented in Table 2. N_1 is the number of experiments, while N_2 is the number of

Table 2: Specified parameters

Parameter	Value
N_1	1000
N_2	5
Initial Stock	15
Season length (weeks)	8
μ_R	$\sim (DiscreteUniform(\frac{1}{5}) - 1) * \frac{1}{2}$
μ_D	$\sim (DiscreteUniform(\frac{1}{5}) - 1) * \frac{1}{2}$
γ_R^m	$\sim (DiscreteUniform(\frac{1}{5}) - 1) * \frac{1}{2}$
γ_D^m	$\sim (DiscreteUniform(\frac{1}{5}) - 1) * \frac{1}{2}$
σ^0	$\sim (DiscreteUniform(\frac{1}{5}) - 1) * \frac{1}{2}$

identical stores simulated at each experiment. We assume that the product is sold in N_2 stores with exactly same demand parameters.

Demand parameters are randomly generated from a discrete uniform function. The values of initial stock and demand parameters are selected in such a way to ensure a balance between demand and inventory on hand. In most of the instances, we see neither a stock out in the first weeks nor an excessive amount of remaining product in the inventory at the end of season. Additionally, trend effect parameters are chosen with the following two constraints 14 and 15.

$$\gamma_D^{tr} > \gamma_R^{tr} > \gamma_R^{or} \tag{14}$$

$$\gamma_D^{tr} > \gamma_D^{or} > \gamma_R^{or} \tag{15}$$

The performance of the algorithm is tested for various cases. First of all, in order to assess the effect of the true product type, half of the simulations is done assuming the true product type is trendy and the other half is done under the assumption that the true product type is ordinary. Then, for each case, we also examine Firm 1's behaviour of decision making. Similarly, for each type of product, half of the instances are simulated assuming Firm 1 is myopic and the other half is generated as if Firm 1 is forward looking. Finally, in order to assess the effect of the discount depth, four discount percentages as 10%, 30%, 50% and 70% are used. As 1000 instances are simulated for each case, in total 16000 instances are generated.

4.2 Effect of trendiness

This section presents the results of the effect of product type on estimation algorithm performance. For each product type 8000 instances examined. In order to obtain consistent results, the parameters other than trendiness parameter are kept same for corresponding instances for each product type. The results are summarized in Figure 3.

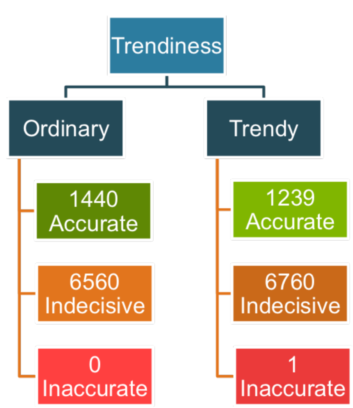

Figure 3: Effect of trendiness

Observation 1. [Accuracy of the method] *Inaccurate results are rarely seen.*

This first observation shows that the algorithm does not mislead. However, there are many indecisive cases where the algorithm could not choose one product type to another. When the parameters representing the product type are close to each other, it is expected that the algorithm cannot simply distinguish the types. In this study, as seen in Table 2, these parameters are deliberately selected close to each other. If they were separate enough, the algorithm would easily detect the types and it would be difficult to assess the real performance of the algorithm under various cases.

Observation 2. [Effect of Trendiness] *Accuracy of detecting the true product type is higher for ordinary products.*

Even though these results do not point out a clear difference between accuracy levels of two types, when true product type is ordinary, the accuracy is slightly higher. It is harder to sell an ordinary product for the firm. The firm tries to increase the demand by offering discounts. Therefore for ordinary products, it is more likely that the demand is observed both under regular and discounted price. This creates an additional information that the algorithm can use to label the product type. For the instances where the trendiness parameters are close to each other, such discounts can also be seen when true product type is trendy. This explains why there is not a distinct difference in number of accurate estimations for each type.

4.3 Effect of decision behaviour

This section tries to answer how Firm 1's decision making process affects the performance of the algorithm. It is examined both assuming the true product type is ordinary and trendy. Here it is assumed that Firm 2 also obtains the information related to Firm 1 behaviour.

Observation 3. [Effect of Decision Behaviour] *Algorithm estimates more accurately when the firm is forward looking.*

Figure 4 summarizes the results where it is clearly seen that the number of accurate estimations are higher when Firm 1 considers the future periods expected profits in decision making process. Not obtaining high accuracy when the firm acts myopically is expected due to the limited information taken consideration by the firm in the value function. When the firm is forward looking, they assess the profits that they are going to make through out the remaining season. The richness of their evaluation enables the model to make better predictions.

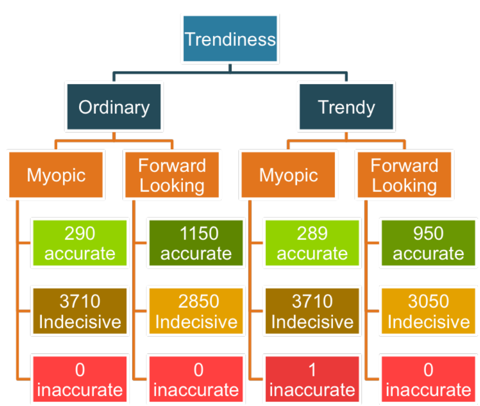

Figure 4: Effect of firm being myopic or dynamic

4.4 Effect of discount depth

Finally, how depth of the discount affects the labelling algorithm is analyzed by dividing instances even more by changing the discount percentages ranging from 10% to 70%. The results are replicated under two product types and myopic and forward looking behaviour of Firm 1.

Observation 4. [Effect of discount depth] *Accuracy decreases as discount depth increases.*

Figure 5 and 6 present the summary of results where it is seen that the increase in discount depth leads to a decrease in the number of successful estimations made by the algorithm. This observation can be explained in a similar way that is pointed out in Section 4.2. As discount depth increases, there is less incentive for the firm to give discounts. The demand increase coming from the discount may not compensate the loss from the price reduction. Hence, for larger discount levels, the algorithm gets less information from sale data and struggles more.

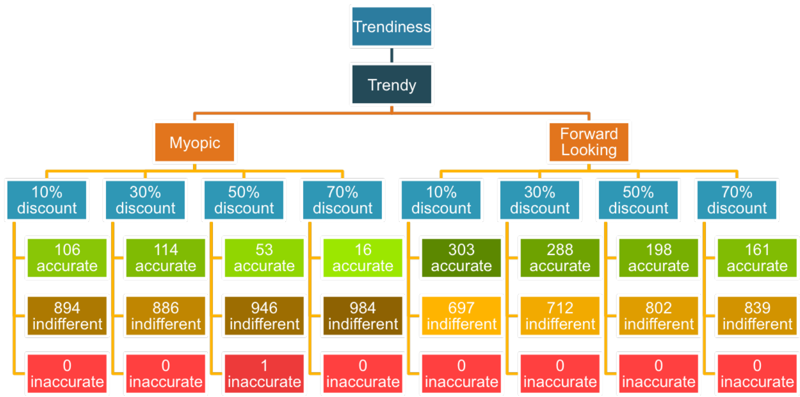

Figure 5: Effect of discount depth when true product type is trendy

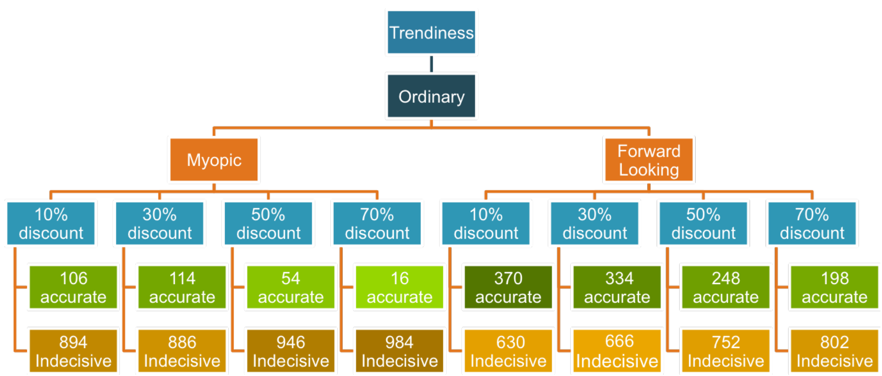

Figure 6: Effect of discount depth when true product type is trendy

5 Concluding remarks and future research directions

This study constructs a labelling algorithm which can be used by a firm who possesses selling data of another firm. The aim of the study is to assess the effect of the each information piece obtained from an outside agent on the performance of estimation. As a preliminary work, discretized cases are studied. For the upcoming works, an extensive approach which includes more levels for price actions could be followed. The product type can be extended in a parameterized way so that the model will not be restricted by two discrete types.

In the future stages of this study, it is also planned to investigate the effect of level of information that Firm 2 obtains. First of all, we would like to examine how the accuracy would change with inclusion of more data from Firm 1. It could be also possible to use less data. How much would the accuracy decrease when the algorithm cannot observe the inventory levels or whether Firm 1 being myopic or forward looking? Would it be possible to depict these informations from pricing decisions?

Finally, we have an extensive dataset taken from a fast fashion retailing company that we would like to apply the algorithm on. It should be noted that the computational complexity of this approach is a limitation.

Since the dynamic programming model is solved exactly in our current algorithm, number of variables in state variables can worsen the computational performance due to the curse of dimensionality. As we plan to implement our algorithm on a large real dataset, the computational issues are needed to be solved. One alternative is to use approximation techniques to solve dynamic program as applied in the literature. We will construct an appropriate approximation approach to overcome the computational problems.

References

Daniel A Ackerberg. A new use of importance sampling to reduce computational burden in simulation estimation. QME, 7(4):343–376, 2009.

Victor Aguirregabiria and Pedro Mira. Swapping the nested fixed point algorithm: A class of estimators for discrete markov decision models. Econometrica, 70(4):1519–1543, 2002.

Victor Aguirregabiria and Pedro Mira. Dynamic discrete choice structural models: A survey. Journal of Econometrics, 156(1):38–67, 2010.

Felipe Caro, Jérémie Gallien, Miguel Díaz, Javier García, José Manuel Corredoira, Marcos Montes, José Antonio Ramos, and Juan Correa. Zara uses operations research to reengineer its global distribution process. Interfaces, 40 (1):71–84, 2010.

Tülin Erdem, Susumu Imai, and Michael P Keane. Brand and quantity choice dynamics under price uncertainty. Quantitative Marketing and Economics, 1(1):5–64, 2003.

Günter J Hitsch. An empirical model of optimal dynamic product launch and exit under demand uncertainty. Marketing Science, 25(1):25–50, 2006.

V Joseph Hotz and Robert A Miller. Conditional choice probabilities and the estimation of dynamic models. The Review of Economic Studies, 60(3):497–529, 1993.

Susumu Imai, Neelam Jain, and Andrew Ching. Bayesian estimation of dynamic discrete choice models. Econometrica, 77(6):1865–1899, 2009.

Masakazu Ishihara and Andrew Ching. Bayesian estimation of finite-horizon discrete choice dynamic programming models. 2016.

Michael P Keane and Kenneth I Wolpin. The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte carlo evidence. The Review of Economics and Statistics, 648–672, 1994.

Michael P Keane and Kenneth I Wolpin. The career decisions of young men. Journal of political Economy, 105(3): 473–522, 1997.

Robert A Miller. Job matching and occupational choice. The Journal of Political Economy, 1086–1120, 1984.

Ariel Pakes. Patents as options: Some estimates of the value of holding european patent stocks, 1984.

John Rust. Optimal replacement of gmc bus engines: An empirical model of harold zurcher. Econometrica: Journal of the Econometric Society, 999–1033, 1987.

John Rust. Using randomization to break the curse of dimensionality. Econometrica: Journal of the Econometric Society, 487–516, 1997.

Kenneth I Wolpin. An estimable dynamic stochastic model of fertility and child mortality. The Journal of Political Economy, 852–874, 1984.