Robustness of the Controlled Variance Pricing Policy

Wendeline Candel-Vergroesen

Business Mathematics and Informatics paper

VU University Amsterdam

De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands

30 December 2010

Preface

This paper is part of the Master program Business Mathematics and Informatics. The objective of the paper is that the student learns how to do an investigation of a business problem on his own and to present the outcomes of this investigation in a correct manner, on paper as well as during an oral presentation.

Arnoud den Boer, who is doing a PhD thesis about Controlled Variance Pricing had some testing to do on his policy. He would be helped with some extra research. I found this very interesting, so I started this investigation energetically and finished with some interesting results that hopefully will help Arnoud with his further research.
In the mean time I also learned to work with Crystal Ball, I wrote in Latex for the first time and made pdf-documents of my pictures, so I learned a lot more than to write a paper.

I would like to thank Sandjai Bhulai for his patience and availability on short notice and during the holidays. He motivated and inspired me during the whole process of writing this paper.

Abstract

Companies that sell products or services have one objective in common: they need to make an optimal profit. To reach that goal they have, among others, to learn the optimal price to sell their products or services for. The process to learn the optimal price is a balance between exploration and exploitation. If they emphasize too much on instant revenue, they will probably not learn the optimal price in the long run. If they emphasize on the learning part they might make unnecessary costs. It helps to use a pricing policy to find the right balance between learning and income. In this paper, I focus on one recently developed policy: Controlled Variance Pricing.
This policy is an extension of Certainty Equivalent Pricing with a shrinking taboo interval. I focus on the performance of the policy with a linear demand function. The performance is measured in the loss a company experiences on their search for the optimal price. This loss is called regret. I explore the basic possibilities of this policy and find that it works very well for linear demand functions with time independent parameters. The regret becomes larger and has higher volatility when the parameters become time dependent. If only one parameter changes, then the relative regret is acceptable in some cases. When more parameters are time dependent one should be very careful for high regrets. A suggestion how to choose the time settings and what history will be considered for the future settings are given in this paper.

Contents

Chapter 1

Introduction

The main goal for a manager in a company that sells products or services is to maximize his profits. This is, among others, done by lowering the costs where possible and by determining the optimal price to reach maximal revenue. To determine the best price he needs to know how the demand behaves for the different prices. If the demand is known, he can calculate the optimal price. The problem in practice is that the demand behavior is unknown. Experimenting with the price is an option to learn the optimal price, but this might be very costly. The manager will be helped with a good policy that balances the loss of revenue and the gain of learning the optimal price to reach his goal of optimal profit. This problem is known as a dynamic pricing problem with uncertain demand while leaving inventory considerations out. Some known policies are:

- Certainty Equivalent Pricing
- MLE-cycle
- One-step ahead pricing
- Controlled Variance Pricing [1]

In this paper I will focus on the last policy: Controlled Variance Pricing.
Controlled Variance Pricing (CVP) is a policy that combines the techniques of Certainty Equivalent Pricing with a shrinking taboo interval. Certainty equivalent pricing sets the new price at each period equal to the price that would be optimal if the current parameter estimates were correct. In: Simultaneously Learning and Optimizing using Controlled Variance Pricing[1] it is proved that Controlled Variance Pricing will converge to the optimal price with an acceptable loss (regret) if the expected demand is a linear function of the price. In the same article it is proved that Certainty Equivalent Pricing does not converge to the optimal price.

The question answered in this paper is: How robust is the Controlled Variance Pricing policy for linear demand functions. I will experiment with the linear demand function where the parameters are time dependent.

The remainder of this paper is structured as follows: In Chapter 2 the CVP model is explained. In Chapter 3 the basic model is implemented and the assumptions will be stated. In Chapter 4 the results of basic parameter settings will be presented and in Chapter 5 different time dependent linear models will be evaluated. In Chapter 6 I will give a conclusion.

Chapter 2

Controlled Variance Pricing

In many situations the demand level for a chosen price is not known. Companies try to maximize their profit. To reach that goal, they need to know the optimal price to maximize the revenue and so to maximize the profit. The process to learn the optimal price might be costly. During the learning process, they do not have the optimal price and loose money in comparison to the optimal situation. The manager is interested in a model that finds the optimal price in a reasonable time with a minimal loss. This loss is also called regret. The Controlled Variance Pricing is such a policy. Controlled Variance Pricing is a refinement of the Certainty Equivalent Pricing policy.

2.1 Certainty Equivalent Pricing (CEP)

Certainty Equivalent Pricing is also called myopic pricing or passive learning. The idea behind this policy is that you learn from your experience in the past and that you use that knowledge to determine your next price. So, for the next price setting, you use the gathered data of the past period. This period can be a week, a day or even a minute. You calculate the optimal price with the corresponding parameter settings of the past period and make this your new price and settings for the next period.

This method is intuitively appealing, but from the study [1] we learn that this method is not consistent. There is a positive possibility that this method does not find the optimal price. From the same paper [1] we learn that the possible cause is the emphasis Certainty Equivalent Pricing puts on instant revenue maximization and not enough on the learning phase. This process is also known as the optimal balancing between exploration and exploitation. With CEP the balance tips to exploitation and not enough on exploration. You can say that this method gets his conclusion too quickly.

2.2 Controlled Variance Pricing (CVP)

CVP solves the problem of the speed by slowing the process to find the optimal price down. This is done with a shrinking taboo interval [1]. The taboo interval is calculated

around the average previously chosen prices. The method calculates the optimal price and the parameter settings of the last period. If the optimal price lies in the taboo interval, instead of taking this value as the new price, they take a suboptimal value outside the taboo interval. In practice this is the border of the taboo interval closest to the optimal price. Each new time step, this interval shrinks as the amount of data grows. The taboo interval takes care of the speed of the process so it is guaranteed to have the time to gather enough data to find the optimal price. In the next chapter this will be visualized and clarified in a figure (see Figure 3.1 at page 13).
This method is proved to be consistent [1] for a linear demand model. For the remainder of this paper I will focus on linear demand functions and the performance of the Controlled Variance Pricing policy.

Chapter 3

Model and notation

In this chapter, I use the same model as in [1]
Assumptions:

- We consider a monopolist firm that sells a single product. At the beginning of each time period $t \in \{1, 2, \ldots\}$ the firm determines a selling price $p_t \in [p_l, p_h]$. The prices $0 < p_l < p_h$ are the minimum and maximum price that are acceptable to the firm.

- The demand D_t in period t is modeled as a linear function of the selling price p_t, with a random disturbance term ϵ_t:

$$D_t = a_0 + a_1 p_t + \epsilon_t. \tag{3.1}$$

  The parameters $\mathbf{a} = (a_0, a_1)$ are unknown to the firm, and are assumed to be in the set $\mathcal{A} = (a_0, a_1) \in (0, \infty) \times (-\infty, 0)$.

- The disturbance terms $(\epsilon_t)_{t \in \mathbb{N}}$ are independent identically distributed random variables with zero mean and variance σ^2.

- We assume that the marginal costs of the sold product equal zero. If the marginal costs are $c > 0$, then by replacing p by $p - c$ we obtain the same model.The revenue collected during period t is denoted by R_t, and is equal to:

$$R_t = p_t D_t. \tag{3.2}$$

  and

$$R_{\text{opt}} = p_{\text{opt}} d_{\text{opt}}. \tag{3.3}$$

  The price $p_{\text{opt}}(\mathbf{a})$, that maximizes the expected revenue per period $E[R_t(p, \mathbf{a})]$, is equal to

$$p_{\text{opt}} = \arg\max_p p(a_0 + a_1 p) = \frac{a_0}{-2a_1}. \tag{3.4}$$

- We assume that $p_{\text{opt}}(\mathbf{a}) \in [p_l, p_h]$.

- A pricing policy ψ is a method that generates a price $p_t \in [p_l, p_h]$ for each t , based on the previously chosen prices $p_1, p_2, \ldots, p_{t-1}$ and on the demand realizations $d_1, d_2, \ldots, d_{t-1}$. This p_t may be a random variable. The performance of a pricing policy is measured in terms of regret. Regret is the expected revenue loss caused by not using the optimal price p_{opt}. For a pricing policy ψ that generates prices $p_1, p_2, \ldots, p_T$, the regret $\text{Regret}(T, \mathbf{a}, \psi)$ after T time periods is defined as

$$\text{Regret}(T, \mathbf{a}, \psi) = \sum_{t=1}^{T} E[R(p_{\text{opt}}(\mathbf{a}), \mathbf{a}) - R(p_t, \mathbf{a})]. \tag{3.5}$$

- Last but not least it is very useful for the comparison of different policies to calculate the relative regret.

$$\text{Relative Regret} = \frac{\text{Regret}(T, \mathbf{a}, \psi)}{TR_{\text{opt}}} 100\% \tag{3.6}$$

Given a number of periods $T \in \mathbb{N}$, the goal of the firm is to find a policy ψ that minimizes $\text{Regret}(T, \mathbf{a}, \psi)$ and finds the optimal price.

- I estimate the parameters (a_0, a_1) with least square regression. The least-square estimates $(\hat{a}_{0t}, \hat{a}_{1t})$ of the parameters (a_0, a_1), based on historical data from the first t periods, are the minimizers of the mean square error:

$$(\hat{a}_{0t}, \hat{a}_{1t}) = \underset{\alpha_0, \alpha_1}{\text{argmin}} \frac{1}{t} \sum_{i=1}^{t} (d_i - \alpha_0 - \alpha_1 p_i)^2. \tag{3.7}$$

If the initial prices p_1, p_2 are chosen different from each other, then (3.7) has a unique solution for each $t \geq 2$ according to [1].

- The taboo interval is the interval in which it is not allowed to choose the next price. This interval controls the order of magnitude of the sample variance. The minimum distance between the new price p_{t+1} and the average price $\bar{p}_t$ is given by:

$$c_0 \cdot t^{\alpha - 1} \leq (p_{t+1} - \bar{p}_t)^2, \tag{3.8}$$

for an $\alpha \in (0, 1)$ and a positive constant c_0. The taboo interval(TI) at time t is given by:

$$TI(t) = \left(\bar{p}_t - \sqrt{c_0} \cdot t^{\frac{\alpha-1}{2}}, \bar{p}_t + \sqrt{c_0} \cdot t^{\frac{\alpha-1}{2}} \right) \tag{3.9}$$

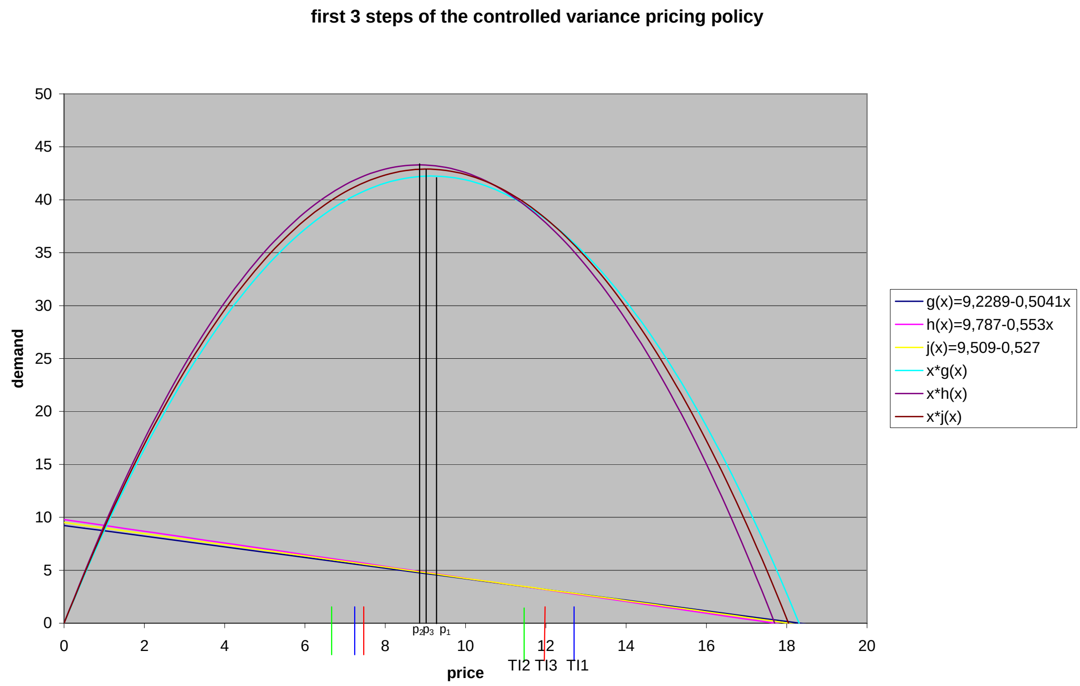

Figure 3.1: example taboo interval

I will visualize and explain the working of the model by Figure 3.1 and Figure 4.2 at page 19. In Figure 3.1 I delineated the first 3 steps of a simulation. I started the simulation with 2 different prices $p_a = 8$ and $p_b = 12$. The program calculates a demand realization with equation $10 - 0.5p_t + \epsilon_t$, where ϵ_t has the standard normal distribution, and the parameters a_0 and a_1 with equation (3.7) (see also the first lines of Figure 4.2). The result is $g(x)$ and with help of the revenue equation (3.2) belonging to these parameter settings p_1 is calculated. The taboo interval(blue lines) is calculated with equation (3.9) with $\alpha = 0.5$ and $c_0 = 10$ and it is determined that p_1 lies in this interval. For the next parameter calculations not p_1 is used but the left bound of the taboo interval, because the left boundary is the closest to p_1. The second step the demand is calculated with the same formula and the parameters are calculated with the three prices and demand realizations. This gives $h(x)$ and the green taboo boundaries. The boundaries depend on the mean of the prices, so we can clearly see the shift of the boundaries. Again lies p_2 in the taboo interval and will a boundary be used instead of p_2. In the third step, with use of all the data so far gathered, $j(x)$ is the result and the red boundaries belong to the taboo interval.

We can clearly see that the taboo interval has shrunken.

Chapter 4

How does CVP behave in basic settings?

To be able to compare different linear demand models and to get a feeling how the policy works, I first implemented the basic model in Crystal Ball. For the taboo interval (3.9) it is needed to choose α and c_0. For all experiments I used $\alpha = 0.5$ since it is proved in [1] that that value is optimal. The effect of c_0 will be examined in my second experiment. In most cases I executed 1000 simulations. I also implemented the Certainty Equivalent Pricing (CEP) policy. In this first experiment we compare CEP with CVP by computing the optimal price for different time settings and comparing the relative regret.

4.1 Comparing Certainty Equivalent Pricing and Controlled Variance Pricing

To start with, I used the demand function $d_t = 10 - \frac{1}{2}p_t + \epsilon_t$, where ϵ_t has the standard normal distribution. $p_l = 5$, $p_h = 10$. The optimal price is according to (3.4) on page 11 $p_{\text{opt}} = 10$. I tested the pricing policies on 4 different time horizons: $T = 25, 100, 500$ and 1000. Because setting the price is still handwork, it does not appeal to use more than a 1000 time frames. If the time scale is an hour it is still more than 41 days that are considered. If the price setting is fully automated it is possible to update the prices each 5 minutes, or even each minute and then time horizons of 10,000 or more are interesting. I chose $c_0 = 10$. The relative regret is computed with the formula (3.6) on page 12. Regret is calculated by the formula (3.5). The regret is computed for each t and after T intervals summed up. The optimal revenue(R_{opt}) is the revenue we can make if we know the optimal price in advance. (see formula (3.3)). In my case this is: $R_{\text{opt}} = T \cdot 50$. The results for the estimated prices and regrets in this basic situation are given in Table 4.1.
We observe that:

- CEP does not converge to the optimal price 10 within the time frames and CVP does.

Table 4.1: Comparing Certainty Equivalent Pricing and Controlled Variance Pricing

T	CEP mean	CVP mean	CEP relative regret	CVP relative regret
25	9.56	9.82	20.34	4.87
100	9.57	9.95	5.81	3.01
500	9.58	9.97	6.33	1.46
1000	9.52	10.00	6.53	0.93

- The accuracy gets better when I take more time intervals for both the prices and the relative regret with CVP.
- The relative regret is lower for CVP.
- CVP works perfect in the situation that the parameters of the linear function are not time dependent.

4.2 How to set the first width of the taboo interval?

I got curious as to the influence of c_0 on the results, so in my next experiment I will try to establish how to choose c_0.
If I look at the taboo interval (3.9) at page 12, I see that the width of the interval equals: $2\sqrt{c_0} \cdot t^{\frac{\alpha-1}{2}}$. We learned already from [1] that $\alpha = 0.5$ is optimal so that reduces the width of the taboo interval to: $2\sqrt{c_0} \cdot t^{\frac{-1}{4}}$. The bigger c_0 the wider the taboo interval. I have to be careful that the taboo interval does not get wider then $p_h - p_l$ otherwise we are not able to chose a new price. So I found that $c_0 \in [0, \frac{\sqrt{t}(p_h-p_l)^2}{4}]$ must apply.
For this experiment I use as demand function $d_t = 10 - 1.5p_t + \epsilon_t$ and $p_l = 1$, $p_h = 6$, I used the time intervals $T = 500$ and 1000. The optimal price according to (3.4) is $p_{\text{opt}} = 3\frac{1}{3}$. Now I focus on finding the influence of c_0 on the estimation of the optimal price and the regret. To compare the regret results with [1] it is even better to calculate the relative regret with (3.6). In this case is $R_{\text{opt}} = 6\frac{2}{3} \cdot T$. I have registered the mean of the estimated optimal price and calculated the relative regret. The results for different c_0 are in Table 4.2.

From these values we notice not many differences, but if we look at the forecast his-

Table 4.2: Results for different c_0 with $d_t = 10 - 1.5p_t + \epsilon_t$

T	$c_0 = 1$	$c_0 = 3$	$c_0 = 5$	$c_0 = 20$
500 (mean)	3.33	3.33	3.32	3.31
1000 (mean)	3.33	3.32	3.33	3.33
500 (rel. regret)	0.092	0.162	0.273	0.023
1000 (rel. regret)	0.461	0.181	0.186	0.039

tograms, we see something very interesting (see Figure4.1 at page 18). The mean value of the prices do not differ much but when $c_0 = 20$ the taboo interval remains relative wide and the realizations of the prices stay far from the optimal price. I compared the width of the taboo interval at the end of the time realizations for $c_0 = 1$ and $c_0 = 20$. In the first case the width of the taboo interval after 1000 time steps is 0.36 and in the second case 1.59. On the other hand, the values of the relative regret with $c_0 = 20$ are very low and if I compare the standard deviations of the mean of the prices or the regret I notice only a small difference compared with smaller values for c_0. The details can be found in the appendix A. I think it worthwhile for further investigation, but for the remainder of my experiments I keep c_0 small.

4.3 What is the influence of the number of data points?

Now I consider the first demand function: $d_t = 10 - 0.5p_t + \epsilon_t$ again but instead of using all data points available so far, I will experiment with the number of data points N that we use. If we can get good results with the use of less data points, that might give some cost reduction opportunities for the manager. The estimation of a_0 and a_1 will be based on N data points but do not forget that the taboo interval also will be calculated on only N data points! We keep $p_h = 15$, $p_l = 5$, $c_0 = 10$ and start with $p_1 = 8$ and $p_2 = 12$.

In Figure 4.2 you see how the implementation in Crystal Ball looks like. In the last column of this picture I see that the taboo interval is really shrinking but that it is not symmetric around one price. This is normal since the taboo interval lies around the mean of the prices known so far. The mean is not constant, so the interval is sliding over the price axes which we also observed in Figure3.1. I also observe that only in 2 cases the calculated p is outside the taboo interval, in all the other cases we use the closest border of the taboo interval. (The calculations are made in a Dutch version of Excel with Crystal ball, so you read Waar and Onwaar instead of True and False).
We use the same time intervals $T = 25, 100, 500$ and 1000. The results for the mean of the price and the relative regret for time intervals: $N = 10, 20, 50$ and 100 are in Table 4.3.

NA means Not Available. Because when $t = 25$, $N = 50$ and $N = 100$ have no meaning.

Table 4.3: Results with limited time intervals for $d_t = 10 - 0.5p_t + \epsilon_t$

T	$N = 10$	$N = 20$	$N = 50$	$N = 100$	$N = 10$	$N = 20$	$N = 50$	$N = 100$
	relative regret				mean price			
25	5.31	4.89	NA	NA	9.87	10.02	NA	NA
100	6.31	4.87	3.40	3.00	9.88	9.84	9.86	9.94
500	6.42	4.91	3.41	2.17	9.80	9.77	9.94	9.91
1000	6.45	4.89	3.40	2.94	9.66	9.73	9.81	9.81

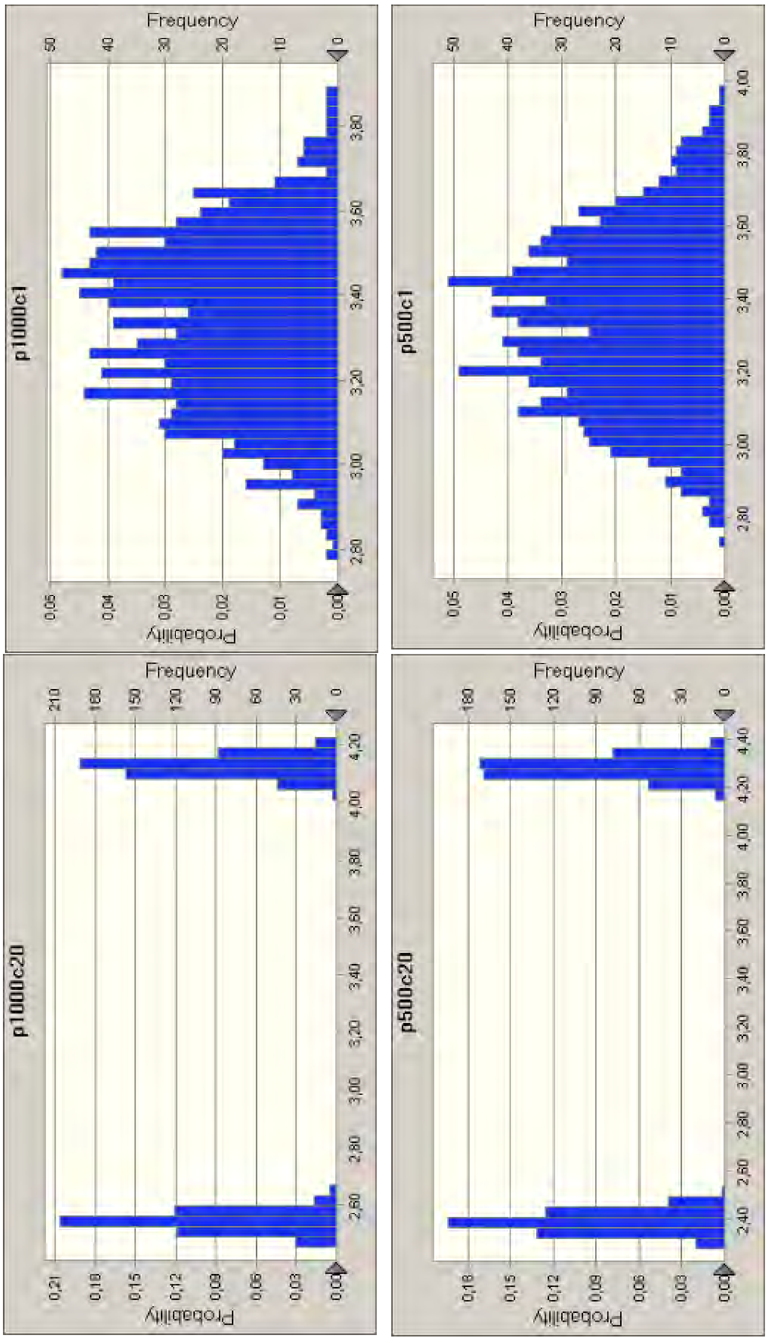

Figure 4.1: Histogram of the forecast of the price for $c_0 = 20$ and $c_0 = 1$

Controlled variance pricing
a1 and a0 will be calculated with linear regression D=ao+a1p+epsilon
over the last prices known in an interval of width N**= 20**
Demand formula:D= 10-0.5p+ε

lowest price	5	p opt	10
highest price	15	D opt	5
c0	10	alpha	0,5

iteration	**price**	**demand**	**a1**	**a0**	**temp price**	**left bound**	**right bound**	**in taboo interval?**	**distance left bound**	**distance right bound**	**bound closest to temp price**	regret	width taboo interval
1	8	5,19626											
2	12	3,17992	-0,50408	9,22893	9,154161	7,34085	12,6591	WAAR	1,813309	3,504987	7,340852	-6,3598	5,318296
3	7,34085	5,87644	-0,5526	9,78702	8,8555	6,71081	11,5164	WAAR	2,144694	2,660929	6,710806	15,626	4,805623
4	6,71081	5,83697	-0,52704	9,50906	9,021164	6,27685	10,749	WAAR	2,744318	1,727818	10,74898	19,199	4,472136
5	10,749	5,16603	-0,41465	8,76648	10,57084	6,84539	11,0749	WAAR	3,725453	0,504032	11,07487	-3,8693	4,229485
6	11,0749	6,2188	-0,2692	7,75266	14,39958	7,29207	11,3331	ONWAAR	7,10751	3,066479	11,3331	-6,6844	4,041031
7	14,3996	3,29134	-0,32352	8,2145	12,69533	8,09517	11,9834	ONWAAR	4,600158	0,711897	11,98343	-14,481	3,888262
8	12,6953	3,10233	-0,36749	8,54482	11,62605	8,491	12,2516	WAAR	3,135045	0,625559	12,2516	-8,3618	3,760603
9	12,2516	3,61543	-0,38017	8,6316	11,3522	8,75448	12,406	WAAR	2,597715	1,053769	12,40597	-8,1405	3,651484
10	12,406	2,83141	-0,41022	8,84661	10,78276	8,98452	12,5411	WAAR	1,798237	1,758321	12,54108	-6,8123	3,556559
11	12,5411	3,63984	-0,41184	8,85863	10,755	9,18805	12,6609	WAAR	1,56695	1,905866	9,188052	-9,2491	3,472816
12	9,18805	5,80421	-0,42973	9,11229	10,60237	9,08072	12,4788	WAAR	1,521653	1,876435	9,080716	4,7127	3,398088
13	9,08072	5,20673	-0,42965	9,11122	10,60302	8,98368	12,3144	WAAR	1,619337	1,71143	8,983681	4,7865	3,330766
14	8,98368	6,84094	-0,4647	9,59377	10,32264	8,8953	12,1649	WAAR	1,427342	1,842283	8,895296	6,9526	3,269625
15	8,8953	5,33416	-0,46205	9,55779	10,34283	8,81426	12,028	WAAR	1,528566	1,685147	8,814264	5,8927	3,213714
16	8,81426	4,93474	-0,451	9,40938	10,43163	8,73955	11,9018	WAAR	1,692076	1,470202	11,90183	5,8513	3,162278
17	11,9018	3,18566	-0,46746	9,53038	10,19386	8,85634	11,9711	WAAR	1,337519	1,777192	8,856345	-6,0586	3,114711
18	8,85634	3,3161	-0,42918	9,01983	10,50826	8,79192	11,8624	WAAR	1,71634	1,354179	11,86244	3,7925	3,07052
19	11,8624	3,41221	-0,43835	9,08807	10,3663	8,89334	11,9226	WAAR	1,472968	1,556327	8,893336	-6,355	3,029295
20	8,89334	6,31346	-0,45757	9,34287	10,20924	8,8369	11,8276	WAAR	1,372334	1,618364	8,836903	6,9869	2,990698

Figure 4.2: Example of a worksheet

- I observe lower optimal prices than in the first experiment, but they are still better than in the CEP case.
- The results for $T = 500$ are better than for $T = 1000$ if we look at the prices and in the case of $N = 100$ is the relative regret for 500 time intervals is below 2.5%. In [1] we learn that a relative regret below 2.5% is a good performance.
- I see that the relative regret improves with the number N used time intervals.
- The relative regret is above 5% for all T and $N = 10$. I think those results high enough to exclude $N = 10$ for my next experiments.
- The relative regret for $T = 100$ and $N = 100$ are the same as in the first example which should be the case, because $N = 100$ with $T = 100$ is the same as without a N.

Now I have enough basic results to start with the experiments with 2 demand functions during the time period.

Chapter 5

Linear demand function with time dependent parameters

In this chapter, I try to find out what the policy does, if after a certain time the demand function changes. I measure only the relative regrets to determine differences between the cases. I set $c_0 = 5$, I keep $\alpha = 0.5$ and I take $T = 1000$ in all experiments. I also take $p_1 = 2$ and $p_2 = 5$ in all cases. I still execute 1000 simulations. I execute my experiments for 3 different cases:

1. Only a_0, the constant, changes. $D_t = 10 - 1.5p_t + \epsilon_t$ and $D_t = 20 - 1.5p_t + \epsilon_t$,
2. Only a_1, the slope, changes. $D_t = 10 - 1.5p_t + \epsilon_t$ and $D_t = 10 - 0.5p_t + \epsilon_t$,
3. The constant and the slope change. $D_t = 10 - 1.5p_t + \epsilon_t$ and $D_t = 20 - 0.5p_t + \epsilon_t$.

I am interested in the influence on the relative regret of this change in demand. I am also curious whether the time at which the change takes place is of great influence. In reality a manager does not know whether the demand function changes and the time when this might happen is also unknown. For a manager it is thus valuable to know that his pricing policy will pick up this change in an acceptable time period and with acceptable regret. I put the change of demand at $T = 50, 250$ and 500 and I use $N = 20, 50$ and 100.

5.1 The constant a_0 changes in the formula $D_t = a_0 + a_1 p_t + \epsilon_t$

The results for case 1: Only a_0, the constant, changes. $D_t = 10 - 1.5p_t + \epsilon_t$ and $D_t = 20 - 1.5p_t + \epsilon_t$ are in Table 5.1. I set $p_h = 15$, $p_l = 1$, $p_{\text{opt1}} = 3\frac{1}{3}$, $p_{\text{opt2}} = 6\frac{2}{3}$.
Note that the optimal revenue is different for the different shock times. For example the optimal revenue for a shock at $T = 50$ in this case is given by the formula: $R_{\text{opt}} = 50 \cdot 3\frac{1}{3}(10 - 1.5 \cdot 3\frac{1}{3}) + 950 \cdot 6\frac{2}{3}(20 - 1.5 \cdot 6\frac{2}{3})$. A visualization of how a_0 changes is made in Figure: 5.1. I see that for all the three values of N, the reaction starts immediately

after the shock, but the larger N, the longer it takes to estimate the new value for a_0. The fluctuation in estimates is lower for higher N.

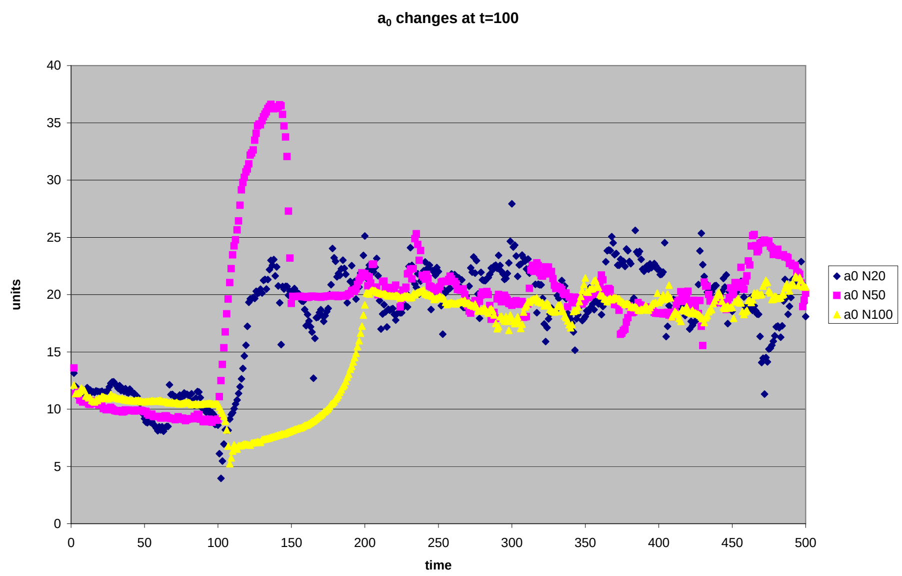

Figure 5.1: A realization of a_0

Table 5.1: Relative regret for changed a_0 in $D_t = 10 - 1.5p_t + \epsilon_t$ and $D_t = 20 - 1.5p_t + \epsilon_t$

Shock at T	$N = 20$	$N = 50$	$N = 100$
50	4.23	3.99	5.92
250	5.18	5.23	6.92
500	6.22	6.84	10.25

5.2 The slope a_1 changes in the formula $D_t = a_0 + a_1 p_t + \epsilon_t$

In the second case: Only a_1, the slope, changes. $D_t = 10 - 1.5p_t + \epsilon_t$ and $D_t = 10 - 0.5p_t + \epsilon_t$ I set $p_h = 15$, $p_l = 1$, $p_{\text{opt1}} = 3\frac{1}{3}$, $p_{\text{opt2}} = 10$. The results are in Table 5.2. I also made a picture of a realization of a_1 for the different N-values. The picture can be seen in Figure

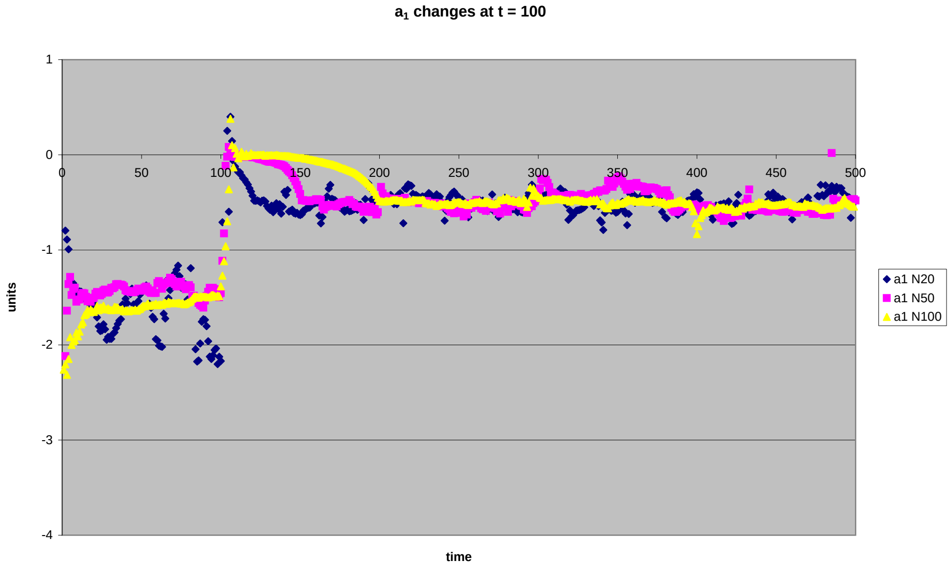

Figure 5.2: Visualization of the development of a_1

5.2. The picture shows that with $N = 20$ the new value is learned faster but the variance stays bigger. It also shows that with $N = 100$ it takes the longest, but then it fluctuates less.

Table 5.2: Relative regret for changed a_1 in $D_t = 10 - 1.5p_t + \epsilon_t$ and $D_t = 10 - 0.5p_t + \epsilon_t$

Shock at T	$N = 20$	$N = 50$	$N = 100$
50	5.99	5.12	4.92
250	6.57	6.95	8.57
500	7.20	7.29	10.89

5.3 The slope and the constant change in the formula $D_t = a_0 + a_1 p_t + \epsilon_t$

In the last case: a_0 and a_1 change. $D_t = 10 - 1.5p_t + \epsilon_t$ and $D_t = 20 - 0.5p_t + \epsilon_t$. I set $p_h = 25$, higher than in the other 2 cases, because the second optimal price is higher,

$p_l = 1$, $p_{\text{opt1}} = 3\frac{1}{3}$, $p_{\text{opt2}} = 20$. These results can be found in Table 5.3.
Now I got in serious trouble with $N = 20$. The system often got stuck because the estimate for a_1 became zero. In Figure 5.3 you can see what happens. This is an example with the shock after $T = 500$, $N = 20$. Because of the shock, the temporary price gets values far outside our assumed interval: $p_t \in [p_l, p_h]$ which is in our case $[1, 25]$. If we get outside our allowed interval, we set the price at the closest boundary. Now we chose 20 times $p_t = 1$ and then our temporary price which is calculated with: $p_{\text{temp}} = \frac{-a_0}{2a_1}$ becomes impossible to calculate because $a_1 = 0$. So my assumption that $p_t \in [p_l, p_h]$ is not correct. As Figure 5.3 shows it will not help to widen our interval, because in this case we got 20 times a negative price which is never allowed. So my advice is not to use $N = 20$ when there is a serious supposition for a variable demand.

Note that the results for $N = 20$ in Table 5.3 are based on only 618 simulations and not

Table 5.3: Relative regret for change of 2 parameters in $D_t = 10 - 1.5p_t + \epsilon_t$ and $D_t = 20 - 0.5p_t + \epsilon_t$

Shock at T	$N = 20$	$N = 50$	$N = 100$	$N = T$
50	4.70	6.24	10.04	**-13.66**
250	5.89	8.10	12.12	**53.35**
500	7.94	11.00	18.57	**43.40**

1000. By accident I also found the following results. I forgot to change the p_h in 25 so it stayed 15 when I simulated the results for $N = 50$ with both parameters changing. The system did not block but I found relative regret values of 39.54% with variance values over 20 million.
It is necessary to look at the variance values of these 3 experiments. They show values between 1 and over 300 million. The mean of the relative regret alone is no more a faithful indicator for the performance of the policy. See for the details of the estimations the Appendix, Section B .
I also worked out the case when I use all the data points. They are in the last column of Table 5.3. Here I can conclude that using all data in this situation is not advisable. It costs a lot and when we look at the variance in Appendix B,Figure B.1. I observe a value of the variance of the mean regret of nearly 719 million and I observe that the simulation broke off at 918 simulations. Findings of my investigation can be summarized as follows:

- I see much larger relative regrets in all the cases.
- I notice that in all the cases it is bad for the relative regret when the shock is at a later time. So the moment the change in demand takes place has a not negligible influence on the relative regret.
- The larger N, the larger the relative regret, with one exception: In the case that only a_1 changes and the shock takes place at $T = 50$, then the relative regret decreases with increasing N.

- I find the standard deviations very high in all the three experiments. I will suggest further investigation of the risks involved with regard to the risks one takes with the regret.
- Do not use all the values so far, because of the high relative regret and the large variability.
- Use at least 50 data points and a wide interval for $p_h - p_l$ to prevent stagnation in the policy.

iteration	price	demand	a1	a0	temp price	left bound	right bound	in taboo interval?	distance left bound	distance right bound	bound closest to temp price	regret
521	2,83252	20,86463	1,129939	16,44161	-7,275439	1,791781	2,727845	ONWAAR	9,06722	10,00328	1,791781	358,1932
522	1	18,76114	0,805432	17,2314	-10,697	1,689986	2,625601	ONWAAR	12,38699	13,3226	1,689986	356,4617
523	1	17,83036	0,616435	17,53813	-14,22546	1,6144	2,549568	ONWAAR	15,83986	16,77502	1,6144	338,7769
524	1	19,04042	0,546862	17,82839	-16,30063	1,553158	2,48788	ONWAAR	17,85379	18,78851	1,553158	361,768
525	1	18,91991	0,479378	18,01664	-18,79168	1,494241	2,428517	ONWAAR	20,28592	21,22019	1,494241	359,4783
526	1	19,37248	0,377796	18,28003	-24,19296	1,436295	2,370127	ONWAAR	25,62926	26,56309	1,436295	368,077
527	1	18,46302	0,387723	18,23894	-23,52056	1,37951	2,312898	ONWAAR	24,90007	25,83346	1,37951	350,7973
528	1	19,3144	0,338183	18,38335	-27,1796	1,323362	2,256308	ONWAAR	28,50296	29,43591	1,323362	366,9737
529	1	18,20767	0,482966	18,18502	-18,82641	1,267659	2,200164	ONWAAR	20,09407	21,02658	1,267659	345,9458
530	1	19,7662	0,352602	18,42949	-26,13358	1,212177	2,144242	ONWAAR	27,34575	28,27782	1,212177	375,5577
531	1	18,80092	0,407224	18,38003	-22,56749	1,15629	2,087915	ONWAAR	23,72377	24,6554	1,15629	357,2174
532	1	19,58349	0,334481	18,52207	-27,68782	1,099634	2,030821	ONWAAR	28,78746	29,71864	1,099634	372,0863
533	1	18,84017	0,24953	18,60226	-37,27464	1,043313	1,974063	ONWAAR	38,31795	39,2487	1,043313	357,9632
534	1	19,38325	0,067808	18,81669	-138,7497	0,987595	1,91791	ONWAAR	139,7373	140,6676	0,987595	368,2817
535	1	20,79661	0,102818	18,92085	-92,01163	0,931958	1,861838	ONWAAR	92,94359	93,87347	0,931958	395,1357
536	1	17,91626	0,225034	18,72807	-41,61157	0,874483	1,803928	ONWAAR	42,48605	43,4155	0,874483	340,4089
537	1	19,68985	0,409207	18,59382	-22,71936	0,817251	1,746263	ONWAAR	23,53661	24,46562	0,817251	374,1072
538	1	18,42176	0,592506	18,38261	-15,5126	0,762074	1,690654	ONWAAR	16,27467	17,20325	0,762074	350,0135
539	1	20,75786	0,885423	18,20861	-10,28244	0,7112	1,639349	ONWAAR	10,99364	11,92179	0,7112	394,3993
540	1	18,67436	0,880443	18,19015	-10,33012	0,664938	1,592657	ONWAAR	10,99505	11,92277	0,664938	354,8129
541	1	20,04334	0,947035	18,18214	-9,59951	0,623618	1,550908	ONWAAR	10,22313	11,15042	0,623618	380,8234
542	1	19,36087	0	19,14021	#DEEL/0!	0,536569	1,463431	#DEEL/0!	#DEEL/0!	#DEEL/0!	#DEEL/0!	367,8565
543	#DEEL/0!	#DEEL/0!	########	########	########	#DEEL/0!	#DEEL/0!	########	########	#DEEL/0!	########	#DEEL/0!
544	########	########	########	########	########	#DEEL/0!	#DEEL/0!	########	########	#DEEL/0!	########	########

Figure 5.3: An example of trouble.

Chapter 6

Conclusions

A manager needs to find out for which price he can sell his products or services best to reach his objective: optimal profit. He can learn by trial and error but should watch out for high costs. He also needs a balance between exploration and exploitation. Controlled Variance Pricing is a policy that can assist the manager in his difficult search for the optimal price. Controlled Variance Pricing is an extension of Certainty Equivalent Pricing with a shrinking taboo interval. The taboo interval guarantees that you will find the optimal price with a bounded regret. In this paper I explored the boundaries of the policy for linear demand functions. I found out that the policy works perfect for time independent parameter settings. The only thing you have to consider in this situation is the value of a constant c_0 which determines the width of the taboo interval. My experiments showed that moderate values are working fine for CVP like $c_0 \in [1, 5]$.
In the situation that the parameters might be time dependent there are some more questions to ask. How much regret is the learning process worth? What level of relative regret is acceptable? The policy is still worthwhile but do not take all the history in to account, because that will become costly and risk full. Also do not take too few time periods in consideration to prevent to stagnate in the policy. I got results for $N = 50$ and $N = 100$ that are in some cases acceptable. The relative regret is highest when the change in demand occurs later in time, but in practice you do not know whether there will be a change or when this will take place.
To further explore the risks of the CVP policy in the situation with 2 time dependent parameters, I advice to experiment with larger time intervals. I would also explore with the width of the taboo interval by experimenting with the value of c_0. I found some promising results in the small exploration I did of c_0, so I advise to look further in the possibilities.

Bibliography

[1] A. den Boer, *Simultaneously Learning and Optimizing using Controlled Variance Pricing* 2010

Appendix A

c0 statistics

Statistics	p1000c1	regret1000c1	p1000 c3	regret1000c3	p1000c5	regret1000c5	p500c1	regret500c1
Trials	500	500	500	500	500	500	500	500
Mean	3,33	10,59	3,32	39,76	3,33	38,48	3,33	57,83
Median	3,34	15,50	3,17	40,87	3,56	35,61	3,34	74,50
Mode	---	---	---	---	---	---	---	---
Standard Deviation	0,21	822,11	0,31	470,59	0,40	379,92	0,25	447,31
Variance	0,04	675.861,06	0,10	221.458,36	0,16	144.336,71	0,06	200.082,93
Skewness	-0,0152	0,1006	0,0490	0,0363	-0,0083	-0,1432	0,0192	0,0862
Kurtosis	2,30	2,99	1,26	2,65	1,10	3,30	2,16	2,96
Coeff. of Variability	0,0629	77,62	0,0946	11,84	0,1194	9,87	0,0743	7,73
Minimum	2,77	-2.326,78	2,82	-1.179,30	2,72	-1.481,19	2,76	-1.041,08
Maximum	3,90	2.637,72	3,85	1.315,99	3,90	1.157,42	3,97	1.583,41
Range Width	1,13	4.964,50	1,03	2.495,29	1,18	2.638,62	1,22	2.624,49
Mean Std. Error	0,01	36,77	0,01	21,05	0,02	16,99	0,01	20,00

Statistics	p500c3	regret500c3	p500c5	regret500c5	p1000c20	regret1000c20	p500c20	regret500c20
Trials	500	500	500	500	1000	1000	1000	1000
Mean	3,33	23,68	3,32	6,67	3,33	6,57	3,31	1,91
Median	3,32	28,31	3,06	-14,08	4,04	9,51	2,47	2,94
Mode	---	---	---	---	---	---	---	---
Standard Deviation	0,38	273,58	0,40	372,58	0,79	204,57	0,95	126,94
Variance	0,14	74.848,38	0,16	138.816,67	0,63	41.848,45	0,90	16.113,01
Skewness	0,0239	0,0478	0,0381	0,2090	-0,0040	0,2104	0,0517	-0,0495
Kurtosis	1,25	2,91	1,09	3,02	1,01	4,68	1,01	2,88
Coeff. of Variability	0,1127	11,55	0,1190	55,87	0,2383	31,14	0,2862	66,30
Minimum	2,69	-702,03	2,76	-1.083,92	2,45	-570,54	2,27	-513,33
Maximum	4,01	899,72	3,88	1.226,24	4,23	1.403,27	4,40	411,52
Range Width	1,31	1.601,75	1,11	2.310,16	1,78	1.973,81	2,13	924,85
Mean Std. Error	0,02	12,24	0,02	16,66	0,03	6,47	0,03	4,01

Figure A.1: Statistics for different c_0

Appendix B

statistics

Statistics	regret shock250 N20 a0	regret shock50 N20 a0	regret shock 500 n20 a0	regret shock250 N20 a0ena1	regret shock50 n20 a0ena1
Trials	1000	1000	1000	668	665
Mean	2.804,46	2.715,67	2.592,42	9.084,55	8.976,96
Median	2.789,30	2.707,24	2.674,42	9.391,70	9.027,24
Mode	---	---	---	---	---
Standard Deviation	1.230,78	1.314,58	1.107,42	4.363,13	5.121,58
Variance	1.514.823,79	1.728.123,01	1.226.386,35	19.036.885,38	26.230.608,81
Skewness	-0,0205	-0,0334	-0,2079	-0,0285	-0,0241
Kurtosis	2,92	2,78	2,63	2,85	2,46
Coeff. of Variability	0,4389	0,4841	0,4272	0,4803	0,5705
Minimum	-898,74	-960,79	-538,48	-2.800,13	-5.041,62
Maximum	6.864,60	7.050,14	5.371,60	22.131,56	21.802,65
Range Width	7.763,34	8.010,93	5.910,09	24.931,69	26.844,28
Mean Std. Error	38,92	41,57	35,02	168,81	198,61
opt revenue	54166,67042	64166,95242	41666,67417	154166,6704	190833,3341
rel regret	5,17746374	4,232188788	6,221805422	5,8926841	4,7040828

Statistics	regret N=T shock250a1	regret N=T shock50a1	regret N=T shock 500a1	regret shock250 N20 a1	regret shock50 N20 a1
Trials	918	918	918	1000	1000
Mean	22.210,61	-6.600,48	14.467,73	2.735,93	2.892,96
Median	46.728,19	-8.168,86	419,94	2.677,02	2.834,82
Mode	---	---	---	---	---
Standard Deviation	26.813,27	3.455,66	16.229,29	1.323,13	1.421,50
Variance	**718.951.623,25**	**11.941.586,20**	**263.389.952,34**	1.750.665,36	2.020.663,23
Skewness	-0,0645	1,06	0,1173	0,1790	0,1462
Kurtosis	1,01	3,59	1,02	3,03	3,03
Coeff. of Variability	1,21	-0,5235	1,12	0,4836	0,4914
Minimum	-7.441,16	-11.173,42	-3.106,36	-1.415,85	-908,55
Maximum	49.745,62	9.720,93	33.834,33	7.575,48	8.410,54
Range Width	57.186,77	20.894,35	36.940,68	8.991,33	9.319,09
Mean Std. Error	884,97	114,05	535,65	41,84	44,95
Cell Errors	0	0	0	1	0
opt revenue	41666,66667	48333,33333	33333,33333	41666,66667	48333,33333
rel regret	**53,30546502**	**-13,65617443**	**43,40318665**	6,56623031	5,98542694

Figure B.1: Statistics

Statistics	regret shock 500 n20 a0ena1	regret shock250 N50 a0ena1	regret shock50 n50 a0ena1	regret shock 500 n50 a0ena1
Trials	667	1000	1000	1000
Mean	8.601,09	12.477,58	11.919,95	11.916,04
Median	8.969,46	14.451,86	11.925,14	15.080,49
Mode	---	---	---	---
Standard Deviation	3.876,91	9.159,49	9.385,02	8.740,18
Variance	15.030.404,18	83.896.347,08	88.078.563,41	76.390.787,52
Skewness	-0,3145	-0,1110	0,0805	-0,1957
Kurtosis	2,91	1,78	2,06	1,59
Coeff. of Variability	0,4507	0,7341	0,7873	0,7335
Minimum	-2.621,76	-8.200,34	-9.769,77	-5.955,01
Maximum	20.759,76	34.518,72	35.723,63	31.141,04
Range Width	23.381,51	42.719,05	45.493,40	37.096,05
Mean Std. Error	150,11	289,65	296,78	276,39
opt revenue	108333,3408	154166,6704	190833,3341	108333,3408
rel regret	7,9394646	8,09356545	6,24626200	10,99942440

Statistics	regret shock 500 n20 a1	regret shock250 N50 a1	regret shock50 N50 a1	regret shock500 N50 a1
Trials	1000	1000	1000	1000
Mean	2.398,46	2.897,52	2.472,58	2.429,67
Median	2.408,71	2.972,71	2.374,78	2.434,31
Mode	---	---	---	---
Standard Deviation	1.120,41	2.329,73	1.914,95	2.111,97
Variance	1.255.324,80	5.427.660,54	3.667.026,38	4.460.425,21
Skewness	0,0219	,00041	0,2193	0,0470
Kurtosis	2,88	2,29	2,94	1,96
Coeff. of Variability	0,4671	0,8040	0,7745	0,8692
Minimum	-814,38	-2.624,31	-2.922,00	-2.683,76
Maximum	7.010,49	9.225,91	9.152,18	8.013,08
Range Width	7.824,87	11.850,21	12.074,18	10.696,84
Mean Std. Error	35,43	73,67	60,56	66,79
opt revenue	33333,33333	41666,66667	48333,33333	33333,33333
rel regret	7,19538675	6,95405257	5,11569130	7,28899514

Figure B.2: statistics (continued)

Statistics	regret shock250 N100a0ena1	regret shock50 N100a0ena1	regret shock 500 N100a1ena0	regret shock250 N100a0	regret shock50 N100a0
Trials	1000	1000	1000	1000	1000
Mean	18.691,86	19.153,51	20.120,38	3.746,32	3.800,68
Median	12.289,58	26.575,49	31.262,92	2.057,68	5.030,39
Mode	---	---	---	---	---
Standard Deviation	16.980,07	17.613,60	16.699,49	4.389,27	4.464,50
Variance	288.322.701,18	310.238.831,26	278.872.840,92	19.265.662,59	19.931.769,96
Skewness	0,0298	-0,0335	-0,1996	0,0536	-0,0393
Kurtosis	1,25	1,32	1,16	1,31	1,37
Coeff. of Variability	0,9084	0,9196	0,8300	1,17	1,17
Minimum	-8.802,50	-12.224,31	-6.731,02	-4.181,47	-5.110,21
Maximum	52.785,45	53.545,30	42.578,23	12.362,97	11.512,68
Range Width	61.587,94	65.769,61	49.309,26	16.544,44	16.622,89
Mean Std. Error	536,96	556,99	528,08	138,80	141,18

	shock250 a0 en a1N100	shock50 a0ena1 N100	shock 500 a0ena1N100	shock250 N100 a0	shock50N100 a0
opt revenue	154166,6704	190833,3341	108333,3408	54166,67042	64166,95242
rel regret	12,12444988	10,03677134	18,57266113	6,916286756	5,923112783

Statistics	regret shock50 n50 a0	regret shock 500 n50 a0
Trials	1000	1000
Mean	2.558,54	2.849,69
Median	2.496,47	3.572,28
Mode	---	---
Standard Deviation	2.439,41	2.298,10
Variance	5.950.743,21	5.281.260,60
Skewness	0,0414	-0,2976
Kurtosis	2,06	1,80
Coeff. of Variability	0,9534	0,8064
Minimum	-3.182,02	-2.410,66
Maximum	9.010,23	8.117,90
Range Width	12.192,25	10.528,55
Mean Std. Error	77,14	72,67

	shock 50 N50 a0	shock 500 N50 a0
opt revenue	64166,66667	41666,66667
rel regret	3,987328899	6,839254251

Figure B.3: statistics (continued)

Statistics	**regret shock 500 N100a0**	**regret shock250 N100 a1**	**regret shock50 N100 a1**	**regret shock 500 N100 a1**	**shock250 N50 a0**
Trials	1000	1000	1000	1000	1000
Mean	4.272,72	3.569,35	2.377,08	3.629,57	2.834,51
Median	6.933,67	2.978,27	2.088,57	5.247,39	3.195,76
Mode	---	---	---	---	---
Standard Deviation	4.406,44	3.703,79	2.532,01	3.663,89	2.410,25
Variance	19.416.673,43	13.718.045,61	6.411.051,84	13.424.101,48	5.809.315,58
Skewness	-0,1695	0,0928	0,3668	-0,0681	-0,0910
Kurtosis	1,21	1,70	2,90	1,33	1,91
Coeff. of Variability	1,03	1,04	1,07	1,01	0,8503
Minimum	-2.950,15	-4.538,31	-4.473,48	-2.842,22	-2.679,50
Maximum	11.609,16	12.791,94	10.705,11	10.253,88	8.819,91
Range Width	14.559,31	17.330,25	15.178,59	13.096,10	11.499,41
Mean Std. Error	139,34	117,12	80,07	115,86	76,22

	shock500 N100 a0	shock 250 n100 a1	shock50 N100 a1	shock 500 N100 a1	shock250 n50 a0
opt revenue	41666,67417	41666,67042	48333,33408	33333,34083	54166,66667
rel regret	10,25453583	8,566435662	4,918105161	10,88869356	5,232949002

Figure B.4: statistics (final)