The University of Chicago Booth School of Business

Course: 37202-01 Pricing Strategy

Pricing Strategies – Final Project

Drew Ficken

Student ID: 12179201

December 12th, 2018

I pledge my honor that I have not violated the Honor Code during the preparation of this assignment.

Drew Ficken, December 12, 2018

Final Project - Fall 2018 - Section 01

Drew Ficken

12/11/2018

```
demand_data <- read.csv("demand.csv")

conjoint_data <- read.csv("conjoint.csv")
```

Phase I tasks:

Question 1:

Based on the survey results, recommend a percentage mark-up for the subscription pricing. Please show your work.

Because our demand curve is derived from just two data points, we just treat the line between them as the demand curve.

```
demand_quantity = c(100, 175)
demand_price = c(400, 250)

pct_delta_quant = (demand_quantity[2] - demand_quantity[1])/(demand_quantity[1])
pct_delta_price = (demand_price[2] - demand_price[1])/(demand_price[1])

Beta = pct_delta_quant/pct_delta_price
# optimal_price = (Beta)/(1+Beta)*cost

optimal_markup = (Beta)/(1 + Beta)

print(optimal_markup)
```

[1] 2

In this scenario we are assuming monopoly pricing and elasticity, so we determine the optimal markup percentage should be 100% more than the cost, or optimal price is 2*cost.

Question 2

What concerns, if any, do you have with this survey and its ability to provide meaningful information for the pricing of the subscription plan? Explain.

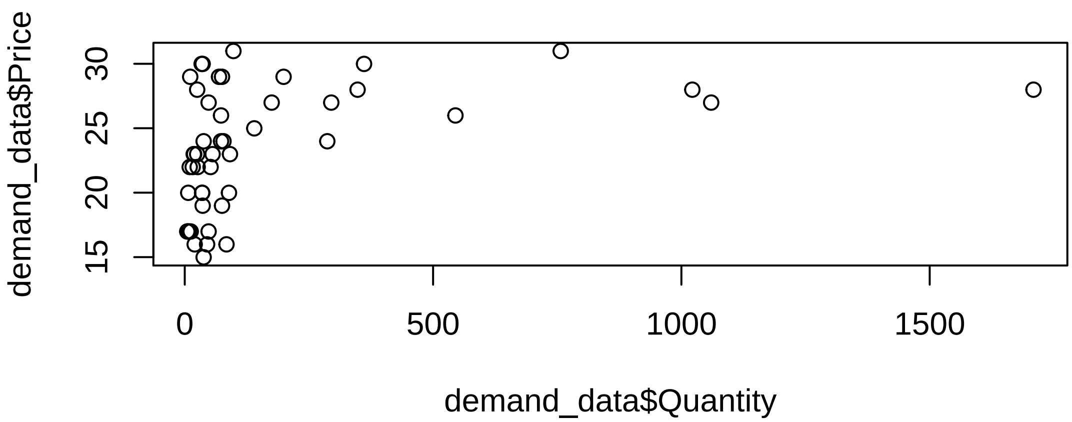
There are a number of issues with this survey. First, there needs to be way more than just two different price points. It would have been much more useful to split up the survey into 10 different price points of 100 participants, rather than 2 price points of 500 participants. This would help to generate an actual curve in the demand curve, rather than a straight downward-sloping line that is just an oversimplification of consumer demand. It would also be important to run the pricing test at different points in the year to look for seasonality. One upside is the survey was conducted at malls across America, (if survye data is actually more granular than the data provided in the above paragraph) so we should be able to look at regional affects, possibly showing local competitive pricing affects. The most important issue by far though is we need much greater variation in prices.

Phase II tasks:

Question 1

Using the demand data in Demand.csv, estimate a log-log demand curve for the existing subscription service. Use whichever software you see fit (e.g. R, Excel etc). Specifically, regress log quantity on log price. Write the mathematical equation for this model. In a well-labeled table, report your regression estimates. Also be sure to report standard errors and tstatistics for each parameter as well as an R-square for the regression. Interpret your parameter estimates. What concerns (if any) do you have with this regression? If you had to recommend a price, what would it be and why? What concerns (if any) would you have with this recommendation?

```
logPrice = log(demand_data$Price)
logQuantity = log(demand_data$Quantity)
plot(demand_data$Quantity, demand_data$Price)
```




```
new_demand_data <- cbind.data.frame(logPrice, logQuantity, demand_data)

reg_1 <- lm(logQuantity ~ logPrice, data = new_demand_data)

r_sq_df <- cbind.data.frame(summary(reg_1)$r.squared, summary(reg_1)$adj.r.squared)
colnames(r_sq_df) <- c("R-squared", "Adjusted R-squared")

kable(tidy(summary(reg_1)))
```

term	estimate	std.error	statistic	p.value
(Intercept)	-7.115422	2.5936067	-2.743447	0.0086377
logPrice	3.577749	0.8249008	4.337187	0.0000781

```
kable(r_sq_df)
```

R-squared	Adjusted R-squared
0.290246	0.2748166

The point of this question is to test whether we just plug data into a regression in R without understanding it, or do we actually understand the output and why we're getting the results that we are. As you can see from the demand curve plot, without providing any sort of filter on the data it appears we have a positive relationship between price and quantity, a.k.a consumers are purchasing more of the product when it gets more expensive. This doesn't follow any sort of logic, so when you dig deeper into the data we are actually seeing extreme seasonal effects, where sales explode during the fall in all four regions (the lowest sales in a fall observation is nearly 4 times the highest sales in another season). While the regression output says that the data is statistically significant leading to a positive slope (p-value < .05 for both intercept and LogQuantity), this is only because the data isn't just random and there is a correlation between price and quantity, it just happens to be positive. Correlation does not equal causation, so we would actually be better off with different data where price is the main contributor to the demand rather than the seasonality if we want to be able to determine an accurate demand curve. If we were to recommend a price based on the data, we would set a price of infinity, but clearly this wouldn't work in the real world so we would have to look to segment the data to reduce/eliminate the seasonality effects.

Question 2

Rerun the above regression while accounting for the possibility of having different base demand for each season. Write the mathematical equation for this model. In a well-labeled table, report your regression estimates. Also be sure to report standard errors and t-statistics for each parameter as well as an R-square for the regression. What concerns (if any) do you have with this regression? If you had to recommend a price, what would it be and why? What concerns (if any) would you have with this recommendation?

Since the data is already formatted as factors for the seasonality and region covariates, we don't need to convert the data at all and can run a regression with the current data.

```
reg_2 <- lm(logQuantity ~ logPrice + Summer + Fall + Winter +
    Spring, data = new_demand_data)

r_sq_df_2 <- cbind.data.frame(summary(reg_2)$r.squared, summary(reg_2)$adj.r.squared)
colnames(r_sq_df_2) <- c("R-squared", "Adjusted R-squared")

kable(tidy(summary(reg_2)))
```

term	estimate	std.error	statistic	p.value
(Intercept)	4.9132043	3.3359126	1.4728216	0.1480818
logPrice	-0.3014861	1.1028574	-0.2733682	0.7858790
Summer	-0.0992518	0.4497062	-0.2207037	0.8263678
Fall	1.9517317	0.5104126	3.8238313	0.0004193
Winter	-1.2876891	0.3716158	-3.4651087	0.0012137

```
kable(r_sq_df_2)
```

R-squared	Adjusted R-squared
0.6476403	0.6148627

Now that we're taking seasonality into account, we do get a negative relationship between logQuantity and logPrice as expected with an estimate of -.30, but the p-value is really high for this term (.7859), making it statistically insignificant. All of the seasons have a significant effect on the intercept of the line we arrive at, with an intercept as low as 4.91 - 1.29, or 3.62 for Winter, and an intercept as high as 4.91 + 1.95, or 6.86 for Fall. This represents a major shift out in the demand curve, so rather than making a pricing decision outright with this data we would want to set a price for each season. This data is still not great for pricing decisions, but it is important in pointing out seasonal effects.

If we did have to set a price, these are the equations we would use for our demand curves.

```
# summer_quantity=exp(reg_2$coefficients[1]+reg_2$coefficients[3])*Price^(reg_2$coefficients[2])
# fall_quantity=exp(reg_2$coefficients[1]+reg_2$coefficients[4])*Price^(reg_2$coefficients[2])
# winter_quantity=exp(reg_2$coefficients[1]+reg_2$coefficients[5])*Price^(reg_2$coefficients[2])
# spring data used for intercept estimate
# spring_quantity=exp(reg_2$coefficients[1])*Price^(reg_2$coefficients[2])
```

But, because we're limited in the amount of data that we're provided with, and the slopes are all the same, the markup should be, if we're following the monopoly pricing and elasticity rule:

```
new_beta = as.numeric(reg_2$coefficients[2])
# optimal_price = new_beta/(1+new_beta)*cost
markup = new_beta/(1 + new_beta)
markup
```

[1] -0.4316108

The issue with this regression is beta is > -1, implying we should be setting a negative price, which is clearly not a good idea in the real world. This is why we need a better data set for pricing strategy. If we are getting a Beta > -1 we are either underpricing our product or we made a mistake with our elasticity measurement, and since the p-value is .7 for our logPrice term it's most likely the latter.

Question 3

Would your conclusions in (2) change if you also included controls for "region" in your regression? Explain why (or why not).

```
reg_3 <- lm(logQuantity ~ logPrice + Winter + Summer + Fall +
    Spring + Region1 + Region2 + Region3 + Region4, data = new_demand_data)

r_sq_df_3 <- cbind.data.frame(summary(reg_3)$r.squared, summary(reg_3)$adj.r.squared)
colnames(r_sq_df_2) <- c("R-squared", "Adjusted R-squared")

kable(tidy(summary(reg_3)))
```

term	estimate	std.error	statistic	p.value
(Intercept)	5.3411576	3.4818196	1.5340133	0.1328988
logPrice	-0.3787184	1.1447104	-0.3308421	0.7424906
Winter	-1.2951919	0.3791839	-3.4157352	0.0014720
Summer	-0.0799949	0.4613872	-0.1733791	0.8632279
Fall	1.9773572	0.5250441	3.7660783	0.0005341
Region1	-0.4220647	0.3640844	-1.1592494	0.2532282
Region2	-0.2251077	0.3631844	-0.6198166	0.5388936
Region3	-0.1328567	0.3645795	-0.3644107	0.7174708

```
kable(r_sq_df_3)
```

summary(reg_3)$r.squared	summary(reg_3)$adj.r.squared
0.6598124	0.6002796

No, including regionality into the regression does not change the conclusion we arrived at in Question 2. The p-value of logPrice is still much higher than .05, making the estimate of -.37 statistically insignificant, and the fact that beta is still greater than -1 implies we are either underpricing our product or we made a mistake with the elasticity measurement, in this case the latter. All of the regions also had p-values greater than .05 so the regions probably shouldn't be included in the regression to begin with as they make the model more complicated and less accurate.

Phase III

Phase III Tasks:

Question 1

Write down the regression model equation for your conjoint analysis and label everything clearly.

The data that we are supplied with shows 15 different subscription profiles that are rated by 1000 different individuals. For the first of the three stages of ex-post segmentation with ratings, we would normally be asked to run a regression separately to obtain subject-specific feature importance weights. Rather than creating a for loop that runs a regression for each subject, we'll just right down what the regression model equation is for each subject.

```
# reg_ind
# <-lm(Rating~NBC+CBS+ABC+HBO+STARZ+CNN+ESPN+Samsung+LG+
# Price, data=conjoint_data[conjoint_data$id==user_input])
```

Question 2

Using the data in conjoint.csv, estimate the relevant regression(s) and report results in a well-labeled table that includes parameter estimates, standard errors and t-statistics for each parameter. Use whichever software you see fit (e.g. R, Excel etc).

Because each individual has the same weight, we can run the regression across all of the data to see which covariates are significant.

```
reg_combined <- lm(Rating ~ NBC + CBS + ABC + HBO + STARZ + CNN +
    ESPN + Samsung + LG + Price, data = conjoint_data)

r_sq_df_4 <- cbind.data.frame(summary(reg_combined)$r.squared,
    summary(reg_combined)$adj.r.squared)
colnames(r_sq_df_4) <- c("R-squared", "Adjusted R-squared")

kable(tidy(summary(reg_combined)))
```

term	estimate	std.error	statistic	p.value
(Intercept)	25.350630	2.0105328	12.608911	0.00e+00
NBC	4.554141	0.9750203	4.670817	3.00e-06
CBS	5.456021	0.6962249	7.836578	0.00e+00
ABC	3.670704	0.8697156	4.220580	2.45e-05
HBO	40.774208	0.5549840	73.469157	0.00e+00
STARZ	11.035803	0.9902433	11.144537	0.00e+00
CNN	4.630804	1.1568239	4.003033	6.28e-05
ESPN	9.157232	1.2148093	7.537999	0.00e+00
Samsung	20.291639	1.1599022	17.494266	0.00e+00
LG	10.544629	1.1076016	9.520237	0.00e+00
Price	-1.108537	0.0548194	-20.221611	0.00e+00

```
kable(r_sq_df_4)
```

R-squared	Adjusted R-squared
0.4859909	0.485648

Question 3

By cluster (i.e. segment), report in a well-labeled table the willingness-to-pay for each attribute. In words, provide a verbal profile of each cluster based on your calculations.

When running through all of the different summaries of the regression outputs, we end up getting extremely high R-squared for all 3 models and all covariates are extremely significant (p-values <.01), so for the table produced I'm just going to show the coefficients stacked up next to each other for each segment.

```
reg_seg_1 <- lm(Rating ~ NBC + CBS + ABC + HBO + STARZ + CNN +
    ESPN + Samsung + LG + Price, data = conjoint_data[conjoint_data$segment ==
    1, ])
reg_seg_2 <- lm(Rating ~ NBC + CBS + ABC + HBO + STARZ + CNN +
    ESPN + Samsung + LG + Price, data = conjoint_data[conjoint_data$segment ==
    2, ])
reg_seg_3 <- lm(Rating ~ NBC + CBS + ABC + HBO + STARZ + CNN +
    ESPN + Samsung + LG + Price, data = conjoint_data[conjoint_data$segment ==
    3, ])

reg_output_segments <- cbind.data.frame(reg_seg_1$coefficients,
    reg_seg_2$coefficients, reg_seg_3$coefficients)
colnames(reg_output_segments) <- c("Segment 1", "Segment 2",
    "Segment 3")

kable(reg_output_segments)
```

	Segment 1	Segment 2	Segment 3
(Intercept)	32.3602657	15.3724439	52.8266458
NBC	7.2201799	0.7347278	15.0744349
CBS	11.3483346	3.5137582	9.9759802
ABC	8.8770890	4.2811384	0.9589751
HBO	40.9725525	52.1052218	8.0806533
STARZ	5.6793926	14.1657617	2.9947352
CNN	1.2283466	5.0365760	4.0837259
ESPN	13.4132705	4.4139339	22.0494117
Samsung	18.2599560	17.0197269	30.0933113
LG	5.2661699	7.5683032	20.0883101
Price	-0.8836557	-1.1546363	-1.0168094

To arrive at the willingness to pay for the different features of each segment, we divide the lambdas by the pricing coefficient (ex: WTP_apple = lambda_apple/ - lambda_price). Doing so we arrive at the following WTP's for each segment.

```
wtp_segment_1 <- reg_output_segments$`Segment 1`/(-1 * reg_output_segments$`Segment 1`[11])
wtp_segment_2 <- reg_output_segments$`Segment 2`/(-1 * reg_output_segments$`Segment 2`[11])
wtp_segment_3 <- reg_output_segments$`Segment 3`/(-1 * reg_output_segments$`Segment 3`[11])

wtp_segments <- cbind.data.frame(wtp_segment_1, wtp_segment_2,
    wtp_segment_3)
colnames(wtp_segments) <- c("Segment 1", "Segment 2", "Segment 3")
row_wtp <- row.names(reg_output_segments)
rownames(wtp_segments) <- row_wtp

seg_size = as.numeric(table(conjoint_data$segment))
segment_percent <- c(seg_size[1]/sum(seg_size), seg_size[2]/sum(seg_size),
    seg_size[3]/sum(seg_size))
```

```
## drop the price row and add the segment size row
wtp_segments <- wtp_segments[1:(nrow(wtp_segments) - 1), ]
wtp_segments <- rbind.data.frame(wtp_segments, segment_percent)
rownames(wtp_segments)[11] <- c("Segment Sizes")

kable(wtp_segments)
```

	Segment 1	Segment 2	Segment 3
(Intercept)	36.620897	13.3136675	51.9533416
NBC	8.170806	0.6363283	14.8252317
CBS	12.842484	3.0431731	9.8110622
ABC	10.045868	3.7077809	0.9431218
HBO	46.367098	45.1269557	7.9470679
STARZ	6.427155	12.2685918	2.9452277
CNN	1.390074	4.3620454	4.0162157
ESPN	15.179294	3.8227915	21.6849016
Samsung	20.664106	14.7403357	29.5958234
LG	5.959527	6.5547074	19.7562201
Segment Sizes	0.045000	0.7090000	0.2460000

What's a little strange about this output is the intercept is basically a willingness-to-pay by each segment of just having the functionality to watch a live feed of television content on their mobile phone device (Nokia), even if they don't have access to a channel.

Segment 2 has the lowest willingness to pay for the base model (Nokia), but they really value the premium channels like HBO and STARZ. It's unlikely they would end up paying for any of the network channels like NBC, CBS, or ABC.

Segment 3 has the highest WTP for the base model (Nokia), but really only values the channels that show sports (NBC, CBS, ESPN). Beyond that they have an extremely high WTP for the Samsung phone or the LG.

Segment 1 is somewhat of an inbetween (though it represents a very small portion of the respondents at only 4.5 percent), with a WTP for the base model of $36.62. They value the network channels more than the other segments, as well as HBO, but beyond that is basically a middle ground between Segments 2 and 3.

Question 4

a)

Consider the current subscription plan offered by the firm and assume it is the only service offered: NBC, CBS, ESPN on the LG device. Plot the EVC demand curve for this subscription plan. Be sure to label your graph clearly.

In this scenario there is no reference product so the reference value is 0. Therefore, the EVC demand curve is just the differentiation value of the three different segments.

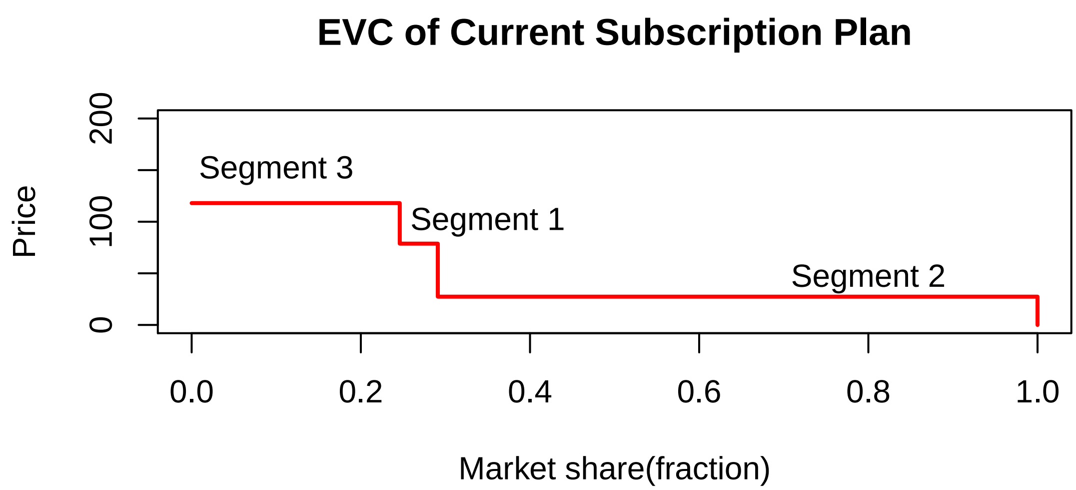
```
EVC.4a = wtp_segments[c("(Intercept)", "NBC", "CBS", "ESPN",
    "LG"), ]

diff_values = colSums(EVC.4a)

evc_graph_points = matrix(c(0, diff_values[3], segment_percent[3],
    diff_values[3], segment_percent[3], diff_values[1], segment_percent[1] +
        segment_percent[3], diff_values[1], segment_percent[1] +
```

```
        segment_percent[3], diff_values[2], 1, diff_values[2],
    1, 0), ncol = 2, byrow = T)

# Now we just make a line plot of the matrix evc3a
plot(evc_graph_points, main = "EVC of Current Subscription Plan",
    type = "l", lwd = 2, col = "red", ylab = "Price", xlab = "Market share(fraction)",
    ylim = c(0, 200)) + text(0.1, 150, labels = "Segment 3") +
    text(0.35, 100, labels = "Segment 1") + text(0.8, 45, labels = "Segment 2")
```



integer(0)

b)

Now suppose the firm also offers an alternative plan consisting of NBC, CBS, ABC, ESPN and STARZ for $50 per month. Plot the new EVC demand curve for the plan offered in 4(a).

Now we have a reference value and the differentiation value changes. We are assuming that the alternative plan is for the Nokia phone since the LG phone is not mentioned explicitly.

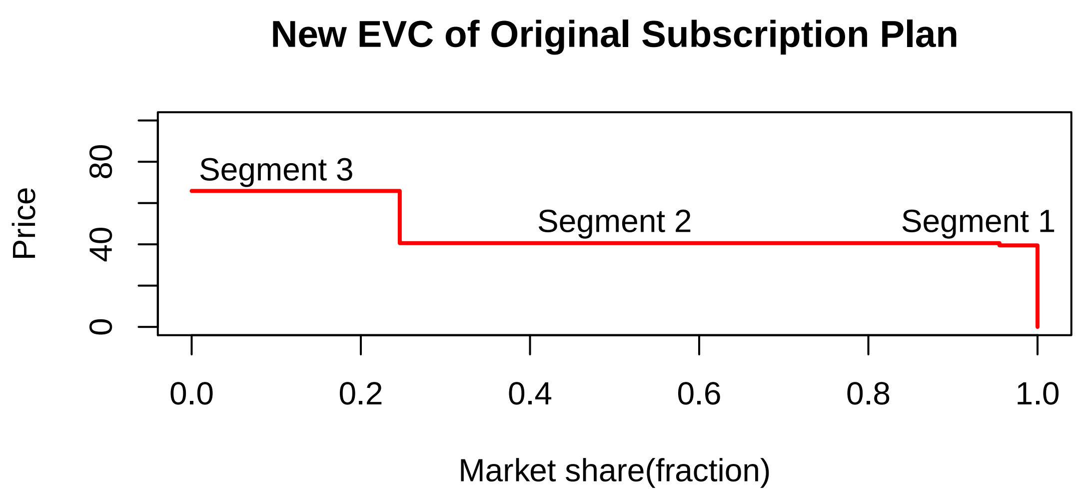
plan 4a = NBC, CBS, ESPN, LG plan 4b = NBC, CBS, ABC, ESPN, STARZ

diff_value = 4a - 4b = + LG - ABC - STARZ

```
ref_value = 50

diff_value_2 = wtp_segments[c("LG", "ABC", "STARZ"), ]
diff_value_2 = rbind.data.frame(diff_value_2[1, ] * 1, diff_value_2[2,
    ] * -1, diff_value_2[3, ] * -1)
diff_value_2_final = colSums(diff_value_2)
evc_values = ref_value + diff_value_2_final
evc_graph_points_2 = matrix(c(0, evc_values[3], segment_percent[3],
    evc_values[3], segment_percent[3], evc_values[2], segment_percent[2] +
        segment_percent[3], evc_values[2], segment_percent[2] +
        segment_percent[3], evc_values[1], 1, evc_values[1],
    1, 0), ncol = 2, byrow = T)
```

```r
plot(evc_graph_points_2, main = "New EVC of Original Subscription Plan",
    type = "l", lwd = 2, col = "red", ylab = "Price", xlab = "Market share(fraction)",
    ylim = c(0, 100)) + text(0.1, 75, labels = "Segment 3") +
    text(0.5, 50, labels = "Segment 2") + text(0.93, 50, labels = "Segment 1")
```




```
integer(0)
```

Question 5

Suppose the firm only offers its current plan, as in 4(a). What is the static, revenuemaximizing, monopoly price of the subscription for a month of service? Show your work.

What we're trying to do is maximize the area under a price point on the graph from part 4a. We'll either use the EVC of segment 1, 2, or 3, depending on what maximizes the revenue. Keeping in mind there were 1000 individuals that took this survey, we arrive at the following conclusion:

```r
seg_3_rev = 1000 * segment_percent[3] * diff_values[3]
seg_1_rev = 1000 * (segment_percent[3] + segment_percent[1]) *
    diff_values[1]
seg_2_rev = 1000 * 1 * diff_values[2]
print(c(seg_3_rev, seg_1_rev, seg_2_rev))
```

Segment 3 Segment 1 Segment 2 29035.57 22922.95 27370.67 The static, revenue-maximizing price is the EVC of segment 3, or $118.03, which allows us to extract all of the willingness-to-pay of that market segment. Even though this price means we don't sell to any of segment 1 or 2, we still generate the most revenue at this price.

Question 6

a)

What are the optimal prices of the phone with basic service and the optional content bundle respectively? Show your work.

We have to set a low enough price on the basic service if we hope to generate revenue from the optional content of the segments that are willing to pay that basic service. Segment 2 clearly has the lowest willingness to pay for both the basic package and the bundle, but they also represent the majority of the market at 70%. We'll try revenues using the EVC's of all 3 segments.

```
phone_val = wtp_segments[c("(Intercept)", "LG"), ]
phone_val_total = colSums(phone_val)
content_bundle = wtp_segments[c("NBC", "CBS", "ESPN"), ]
content_bundle_total = colSums(content_bundle)

rev_content_1 <- 1000 * (segment_percent[1] + segment_percent[3]) *
    (as.numeric((content_bundle_total[1])))

rev_content_2 <- 1000 * (1) * (as.numeric((content_bundle_total[2])))

rev_content_3 <- 1000 * (segment_percent[3]) * (as.numeric((content_bundle_total[3])))

rev_phone_1 <- 1000 * (segment_percent[1] + segment_percent[3]) *
    (as.numeric((phone_val_total[1])))

rev_phone_2 <- 1000 * (1) * (as.numeric((phone_val_total[2])))

rev_phone_3 <- 1000 * (segment_percent[3]) * (as.numeric((phone_val_total[3])))

content_rev <- cbind.data.frame(rev_content_1, rev_content_2,
    rev_content_3)
colnames(content_rev) <- c("Segment 1", "Segment 2", "Segment 3")

phone_rev <- cbind.data.frame(rev_phone_1, rev_phone_2, rev_phone_3)
colnames(phone_rev) <- c("Segment 1", "Segment 2", "Segment 3")

rev_output <- rbind.data.frame(phone_rev, content_rev)
rownames(rev_output) <- c("basic package rev", "optional content rev")

kable(as.data.frame(content_bundle_total))
```

	content_bundle_total
Segment 1	36.192584
Segment 2	7.502293
Segment 3	46.321196

```
kable(as.data.frame(phone_val_total))
```

	phone_val_total
Segment 1	42.58042
Segment 2	19.86837
Segment 3	71.70956

```
kable(rev_output)
```

	Segment 1	Segment 2	Segment 3
basic package rev	12390.90	19868.375	17640.55
optional content rev	10532.04	7502.293	11395.01

```
total_revenue = phone_rev[2] + content_rev[3]
```

The optimal prices of the basic service and the optional content are $19.87 and $46.32.

b)

How do the revenues and prices from this scheme compare with the scheme in Q5? Does your answer make intuitive sense? Why or why not?

The revenues increase slightly because the optimal price for the phone is the lowest price, which allows us to capture the entire market. The price of the optional content remains the same, where in terms of pricing we only hope to capture Segment 3's WTP. This makes sense because Segment 2 wasn't willing to pay for anything in the bundled scenario, but now that we can purchase these two things separately the overall revenue from the low phone price is higher ($19,868) than the revenue where we only capture segment 3 ($17,640). Overall increase in revenue is then just the difference between these basic package's revenues. This makes intuitive sense as well since we're tapping into a previously unused source of revenue. If the revenue from Segment 2 on the phone package had been less than that of Segment 3, then there would be no difference between the answer to Question 5 or Question 6a. Overall revenue in the unbundled package is $31,263, whereas bundled package is $29,035.

Question 7

a)

What fraction of the prospective consumer market would you predict will adopt the service at this price? This fraction constitutes the probability that a randomly selected prospective customer would adopt at the given price level.

We assume that the segmentation of the test market (1000 individuals surveyed) is a representation of the entire market, so the breakdown should be the same. This is the segment percent of segment 3, which is 24.6%

b)

Build a LTV model and calculate the net present value of a prospective and an existing customer respectively. Assume that the customers are homogeneous in their taste for the service and a prospective customer is contacted once per month. Show your work.

Current Customer

LTV of a current customer is the NPV of all future revenues from a customer.

Monthly price = $118.03 (from question 5) Monthly cost = $8 Monthly revenue = $118.03 Monthly discount rate = 1/(1 + .05) = .952 Monthly retention rate = 30% LTV = monthly revenue/ (1 - monthly discount rate * monthly retention rate)

Potential Customer

LTV of a prospective new customer = Acquisition rate * (LTV of an acquired customer) Acquisition rate = 24.6%

```
LTV_current = 118.03/(1 - 0.952 * 0.3)
LTV_prospective = (LTV_current) * 0.246
print(c(LTV_current, LTV_prospective))
```

[1] 165.21557 40.64303

c)

Recently, HBO has contacted your firm to negotiate a contract that would add HBO as part of the monthly subscription plan. The proposed contract would bill your firm an upfront fixed payment of $50 million irrespective of the number of subscribers and a monthly payment of $2 per subscriber. Suppose you would continue to charge the price calculated in question 5 even after adding HBO to the subscription. Would adding HBO to the subscription package be profitable in the long term? Show your work.

We need to recalculate the EVC's to see if we acquire more customers at this price point.

```
EVC.7c = wtp_segments[c("(Intercept)", "NBC", "CBS", "HBO", "ESPN",
    "LG"), ]
diff_values_7c = colSums(EVC.7c)
```

We now have two different segments that are willing to pay the price of /$118.03, but the segment 1 size that is added is relatively negligible (only 4.5% of the prospective market). We need to look at the overall revenue generated from the addition of this segment, meaning that acquisition rate would increase from 24.6% to 29.1%.

The $2 fee only affects current customer LTV, and the added acquisition rate only affects prospective customer LTV. We need to look at the difference in revenue generated, knowing the other costs and revenue's aren't going to change. If the increase in revenue is >50 million, then adding HBO to the subscription package is profitable long term

```
change_current_LTV = -2/(1 - 0.952 * 0.3)
new_prospective_LTV = 118.03/(1 - 0.952 * 0.3) * (0.246 + 0.045)
change_prospective_LTV = new_prospective_LTV - LTV_prospective
current_customer_base = 150000
prospective_customer_base = 5e+05 - current_customer_base
profit_change = current_customer_base * change_current_LTV +
    prospective_customer_base * change_prospective_LTV
profit_change
```

[1] 2182212

Profit only increases by 2.18 million, much less than the 50 million dollar fee HBO was going to charge. Keep in mind that this is under the assumption that the above equation assumes that the monthly retention rate remains at 30%, something that would most likely change given that the product is better for the same cost. It would make more sense to go back and determine a new optimized price now that the product that we are offering has changed, possibly dropping it below the EVC of $72.49 to get the other 72% of the untapped market in segment 2.