DRAFT VERSION DECEMBER 11, 2016
Preprint typeset using LaTeX style AASTeX6 v. 1.0

WHICH PROMOTIONS SHOULD WE BE OFFERING?

VISHAL KASLIWAL[1,2]

[1]Department of Physics & Astronomy, University of Pennsylvania
[2]Department of Astrophysical Sciences, Princeton University

ABSTRACT

CubeSmart allows customers to reserve storage units at no charge. Only a fraction of reservations turn into rentals i.e. some customers are 'no shows'. CubeSmart offers a number of promotional offers as incentives for customers to actually rent reserved storage units. All promotions only apply for a specified time period, usually the first one to two months. Such promotions can be categorized into two broad groups—promotions that offer rent-free storage, and promotion that offer a reduced rental rate. We examine a dataset consisting of reservation records with the goal of quantifying the benefits of the promotion program to CubeSmart in the form of revenue. We find that reduced rental rate promotions can potentially generate $\sim 1.5\times$ the revenue generated in the absence of all promotions. Promotions that offer rent-free months generate reduced revenue compared to promotion-free rentals. An examination of the duration for which the units are rented suggests rent-free promotions probably attract an over-abundance of customers with short-term storage needs who end up generating little revenue. We discuss potential strategies to convert such customers to revenue-generating customers. We conclude by discussing potential improvements to this study.

1. APPROACH

The most interesting quantity in this exercise is the revenue generated under the various promotions as compared to the revenue generated in the absence of all promotions. Including the absence of any promotions, the promotions are -

- No Promotion (NP)
- First Month Free (FMF)
- First Month Half-Off (FMHO)
- Two Months Free (TwMF)
- Two Months Half-Off (TwMHO)
- Three Months Half-Off (ThMHO)

We find that there seem to be no incidences of people renting with the Three Months Half-Off promotion and so we will not consider it for the remainder of this discussion.

The dataset spans a time period from 18:31:57 on 11/10/2012 to 18:43:47 on 06/11/2014 and provides several items of interest—we list the most interesting quantities. We also include some quantities that are computed using the provided data (italicized). These quantities are -

1. Account ID: Unique reservation record identifier.
2. Store ID: Unique storage location identifier for the store at which the unit was reserved.
3. Unit Size: Size in square feet of the reserved storage unit. Units with size 0 are storage spots for motor vehicles and are discussed in detail later.
4. Rented?: Boolean that identifies if reservation $\rightarrow$ rental.

vishal.kasliwal@gmail.com

5. Reservation Date: The date & time of the reservation.
6. Move In Date: The date & time of the rental. This field is the exact same as the reservation date if the customer did not rent the unit.
7. Move Out Date: The date & time of at which the rental ended. This field is left unpopulated if customer did not rent the unit or if unit had not been vacated teh unit as of the end of the data-collection period.
8. *Rental Duration*: The duration for which the customer rented the unit. If the customer had not vacated the unit as of the end of the data-collection period, then this field is populated with the total time since the rental began to when data-collection ended.
9. Promotion: Only populated when the customer was offered a promotion.
10. Rent Rate: Monthly rate at which the customer was offered the rental.
11. *Rate Per Square Foot*: Computed as the monthly rate divided by the square footage of the reserved unit for storage units and as the monthly rate for motor vehicle spaces.
12. Region: Geographic region of the rental.
13. Survey Responses: Only recorded for customers that rented a reserved unit.

An examination of the dataset reveals that a hand-full of entries are suspect due to incorrect Reservation/Move In/MoveOut Dates. We begin by removing suspect entries, replacing null values with appropriate placeholders, and computing quantities of interest. Refer to the analysis code for details.

Broadly speaking, we are interested in the behavior of the quantity

$$R = p(r_{\rightarrow}|\Theta, \kappa)D(\Theta)\kappa, \tag{1}$$

where κ is the rental rate, $p(r_{\rightarrow}|\Theta, \kappa)$ is the likelihood with which customers rent the storage location that they have previously reserved given the conditions Θ and rental rate κ, and D is the duration of the rental. We are interested in

1. identifying the conditions Θ that affect the rental likelihood $p(r_{\rightarrow}|\Theta)$.
2. quantifying the behavior of the rental likelihood $p(r_{\rightarrow}|\Theta, \kappa)$ as we vary the conditions Θ & the price κ.
3. quantifying the behavior of the rental duration $D(\Theta)$ as we vary the conditions Θ.

with the goal of

- identifying the κ and Θ that maximize R.

In general, many factors could enter into Θ. Ideally we would like to consider the most significant factors. We begin by discarding motor vehicle storage unit reservations from the analysis. Since these types of units are provided at a fixed square footage, it is likely that their pricing is governed by different conditions as compared to storage units. For example, heritage motor vehicles require good climate control to ward off degradation due to the environment whereas ordinary motor vehicles are effectively self-sealing and require no climate-control. Given the large number of motor vehicle storage reservations (~ 3000), we feel that it would be more appropriate to analyze motor vehicle storage spots separately.

We examine the relationship between the rental cost (κ) and the rental duration for storage units that had been rented out and subsequently vacated. As a proxy for the rental cost, we shall use the rental rate per square foot (ρ) in order to analyze all the remaining units simultaneously. The rental rate per square foot is calculated as

$$\rho = \frac{\kappa}{A}, \tag{2}$$

where A is the rental unit area in square foot. Segregation by unit size could possibly prove useful in a more detailed analysis in the future. Figure 1 shows the relationship between the rental rate offered to the customer and the duration that the customer rented the unit for. We color the individual records by promotion type offered and conclude that the rental duration depends weakly on the rental rate offered to the customer. More significantly, the rental duration depends significantly on the promotion type offered as evidenced by the presence of a large number of green (First Month Free) points at short rental durations. We compute the average rental rate per square foot for each promotion type and find

Promotion	ρ ($\$/ft^2$)
No Promotion	1.10
First Month Half Off	1.12
Two Months Half Off	1.19
First Month Free	1.08
Two Months Free	1.04

Next we attempt to characterize the duration that customers rent for $D(\Theta)$ using a *survival analysis*. Survival analyses are appropriate when handling temporal data with censorship i.e. data that quantifies a duration but where some events may not have been observed to have finished. For example, consider an analysis of the lifetimes of the buildings in a city. City records will provide construction and demolition dates for all buildings that are no longer standing. However, we have no estimate of the eventual future lifespan of all the existing buildings in the city since we have not yet observed their demolition. In this case, we are interested in rental durations but our data do not have move out dates for renters who were still renting as of the end of data collection. We use the survival analysis to estimate the median rental duration i.e. the interval of time that has passed after the beginning of the rental period at which we have a 50% chance of finding that the renter is still occupying the unit. We find that for some promotions, we are unable to non-parametrically estimate the survival function to the 50%-level. In such cases, we fit a smoothing spline to the estimated survival function and use the spline to estimate the median rental duration for that promotional type. This method is ad-hoc at best. Additional data or a better choices for the parametric form of the survival function would greatly aid in refining the median rental durations. Figure 2 shows the estimated survival functions for each promotion type. We compute the median rental duration for the different promotion types and find

Promotion	D (d)
No Promotion	529.42
First Month Half Off	318.58
Two Months Half Off	2756.39
First Month Free	342.30
Two Months Free	858.18

Lastly, we characterize the rental likelihood $p(r_{\rightarrow}|\Theta, \kappa)$ for the different promotion types Θ as the rental rate per square foot ρ varies. We find that the best fit regression coefficients are

Coefficient	Value
Intercept	0.774610837246
First Month Half Off	0.546097764356
Two Months Half Off	0.178227289856
First Month Free	0.405176906148
Two Months Free	-1.66899921413
ρ	-0.996248929968

We observe that the rental likelihood is strongly anti-correlated with the rental rate per square foot. We also observe that the most liked offers are First Month Half Off and First Month Free. Figure 3 shows the implied likelihood that reservations will convert into rentals as a function of the rental rate per square foot. We speculate that the Two Month Free likelihood curve is to be suspect because of the small number of reservations that were offered that promotion.

Figure 4 plots the revenue curves inferred from the survival analysis median rental durations given the likelihood model generated by the logistic regression performed earlier. We estimate and plot the rental rate that generates maximum revenue per square foot (dashed lines) for each promotion type along with the average rental rate per square foot charged to customers based on the reservation records (dotted lines). It is notable that the analysis suggests that CubeSmart is generally charging a lower rental rate per square foot than what is predicted to generate maximal revenue.

2. DISCUSSION

We have performed two statistical analyses in the process of generating the revenue curves of Figure 4—a survival analysis & a logistic regression analysis. We have made a number of simplifications and approximations in the interest of time. To make the results more accurate, we that a survival regression analysis be performed of the rental durations against the promotion types, the storage unit sizes (categorical), the actual rental market rate, the presence or absence of the 10% promotional discount, & the move-in month. We speculate that there may exist a strong trend in the rental duration with the rental month e.g. from students looking for summer storage etc... With more data, we should also consider the rental store location (rather than the broader geographic location) to model trends as function of distance from major population centers. Additionally, we could also take the survey response answers into consideration as habitual renters may be moving into/out of CubeSmart locations in order to reduce their rental costs. The survival regression model can be used to predict mean/median rental durations for different combinations of the underlying regressors. Similarly, the regression analysis performed to determine the rental propensity with rental rate should be made more sophisticated by including the added regressors mentioned above.

We can use the regression model and the rental duration estimates computed earlier to determine how much likelier the promotions make reservations $\rightarrow$ rentals as well as how much revenue the promotions are currently generating for CubeSmart at the average rental rate per square foot currently being charged to customers. The current revenue per square foot is

Promotion	$p(r_{\rightarrow}\|\Theta, \kappa_{\text{curr}})$	Revenue ($\$/ft^2$)
No Promotion	0.610481	13.48
First Month Half Off	0.726427	15.09
Two Months Half Off	0.633079	19.53
First Month Free	0.706704	9.15
Two Months Free	0.239965	4.23

In conclusion, we can answer a few interesting questions -

Q. Do the promotions result in more rentals?

A. All the promotions other than the Two Months Free promotion result in a higher rate of rentals as compared to when customers are offered no promotions at all.

Q. Do the promotions result in increased revenue?

A. Only the half-off promotions result in increased revenue for CubeSmart. The free-month/s promotions result in considerably reduced revenue for CubeSmart. This probably due to the nature of the storage requirements (short term as opposed to long term) of the customers who use the free-month promotions.

Q. Based on the ”cost” (in lost revenue) of providing free month promotions, how much of an increase in rental or reservation volume would CubeSmart have to see to ”break even” on the promotion?

A. Lets assume that CubeSmart has to choose between offering 2 promotionless rentals and offering x promotionless less rental $+1$ Free Month rental. Using the revenue values from the revenue table, we see that

$$2 \times 13.48 = 9.15 + x \times 13.48 \implies x = \frac{2 \times 13.48 - 9.15}{13.48} \implies x = 1.321 \text{ rentals},$$

i.e. CubeSmart would have to see a $2.321/2 \sim 16\%$ increase in the number of rentals coming in for it to be possible for CubeSmart to break even on offering the Free Month promotion. Similar values can be easily computed for the other promotions.

Q. Free month promotions are effectively a discount for the customer, though hopefully generate some ultimate benefit to revenue. Using the dataset, can you assess how much incremental revenue is generated (or lost) from the free month promotions?

A. Please look at the values in the likelihood & revenue table for the median lifetime revenue per square foot expected from the rentals with and without the promotions.

Q. To what degree does the length of customers’ stays impact your responses above and how should it be considered? What is the most appropriate way to consider customer length of stay, given the number of customers still renting?

A. Since the rate per square foot for the rental is higher when offered with a promotion, CubeSmart starts turning a larger and larger profit as the rental duration increases. A survival regression analysis will be very helpful in further illuminating the rental duration behavior.

CubeSmart seems to be loosing money by offering the popular First Month Free and the less popular Two Months Free promotions. Both the First Month Half Off and Two Months Half Off promotions generate increased revenue for CubeSmart as compared to not offering any promotions at all. The author suspects that the free month proposals are used by a large number

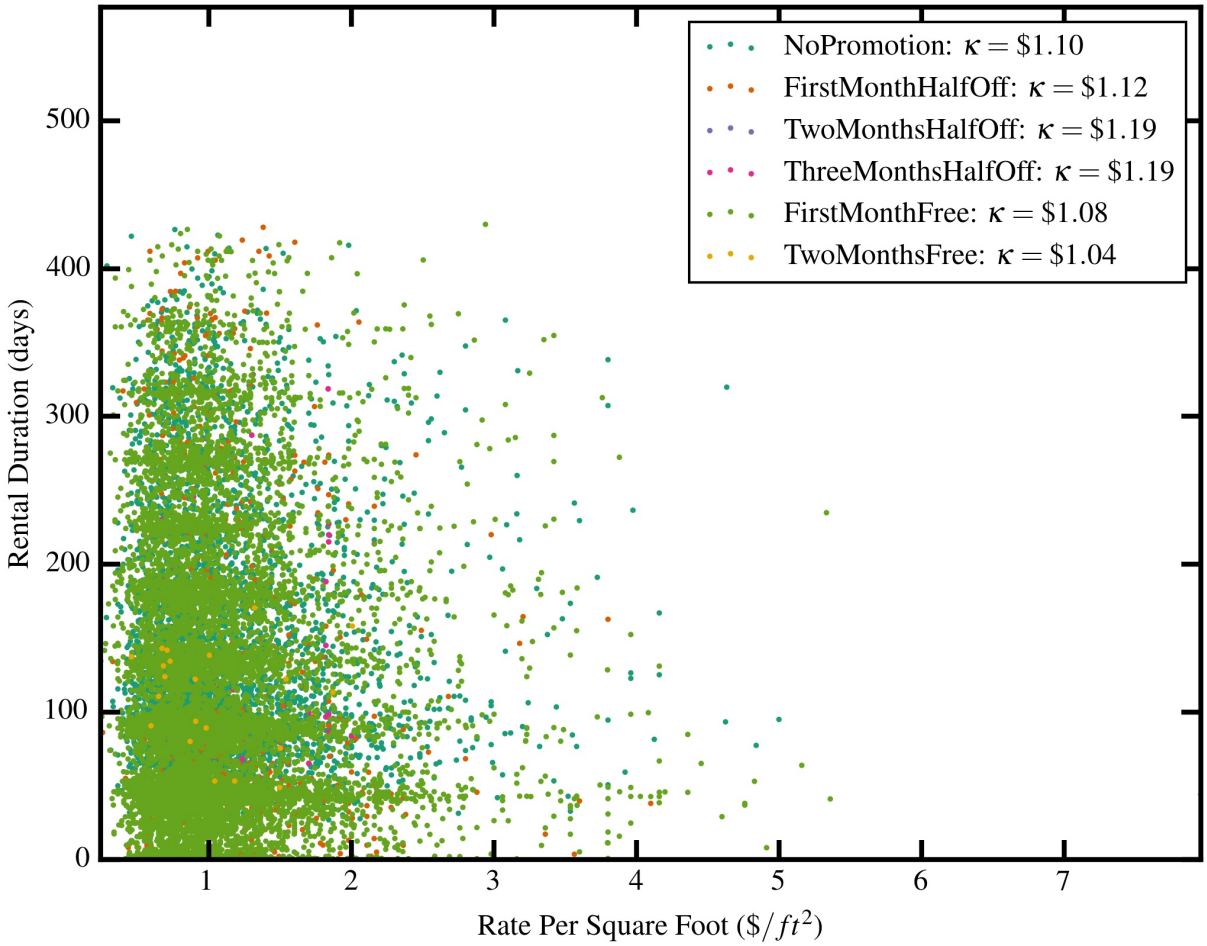

Figure 1. Rental duration as a function of the rental rate per square foot. Each point is a single rental record and is colored by the type of promotion offered. Notice the weak trend in the rental duration as the rental rate per square foot varies. Also significant is the slight clustering in rental duration by promotion type suggesting that the rental duration may depend on the promotion type offered.

of short-term renters who move out before, or soon after the end of the promotion period. In contrast, the half-off promotions are probably attracting regular users of personal storage who rent for a much longer duration at the slightly higher base rate commanded by the reservations tied to these promotions resulting in the increased revenue after the expiration of the promotion period. At the same time, the free month promotions carry real advertising value in terms of familiarizing potential future long-term customers with CubeSmart as the the provider of their storage needs. An alternative strategy may be to offer free storage to customers who move out within a very short duration (say 1 - 2 days) instead of the free month offers.

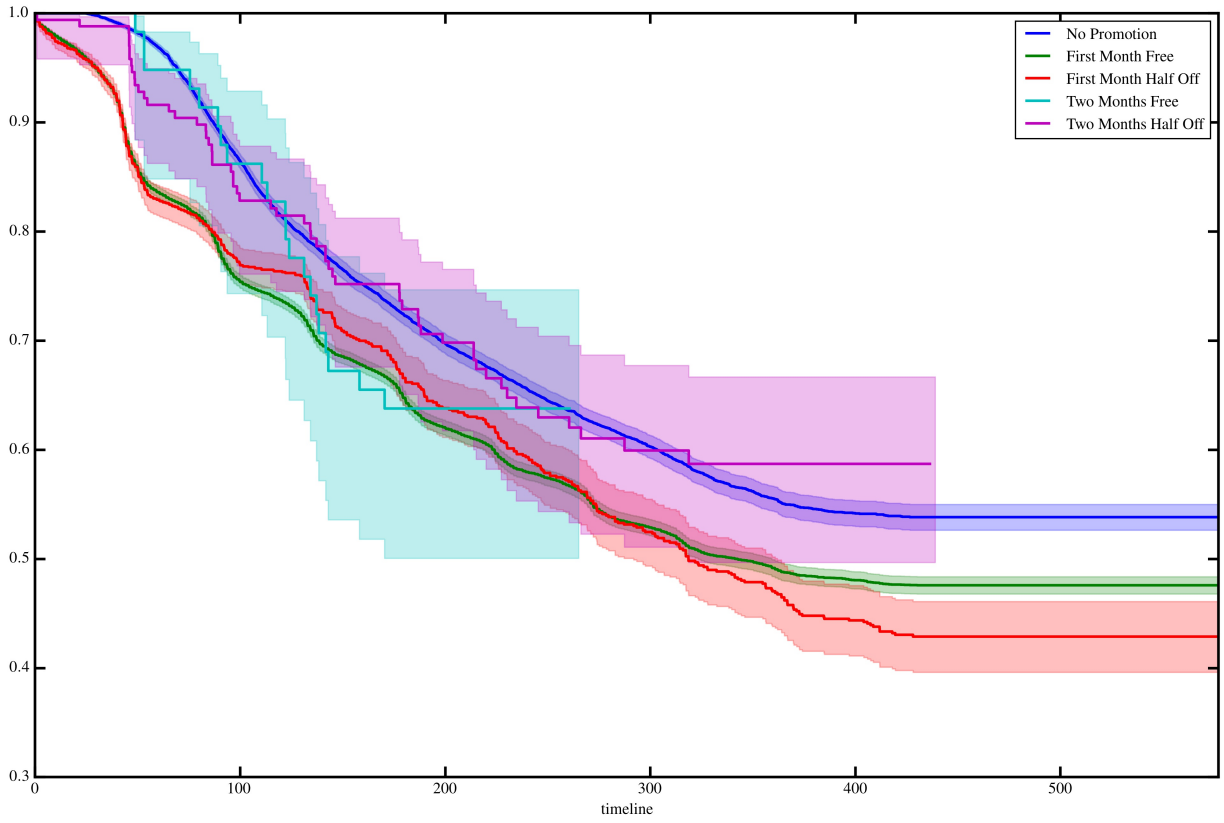

Figure 2. Survival functions for customers renting units with different kinds of promotions. The y-axis plots the likelihood that a customer will no longer be renting a storage unit after x-axis time t. In some cases, we have insufficient data to compute the survival functions down to the 50% likelihood level. In such cases, we use a smoothing spline to estimate the rental duration at which there is a 50% chance that the customer is still renting with CubeSmart.

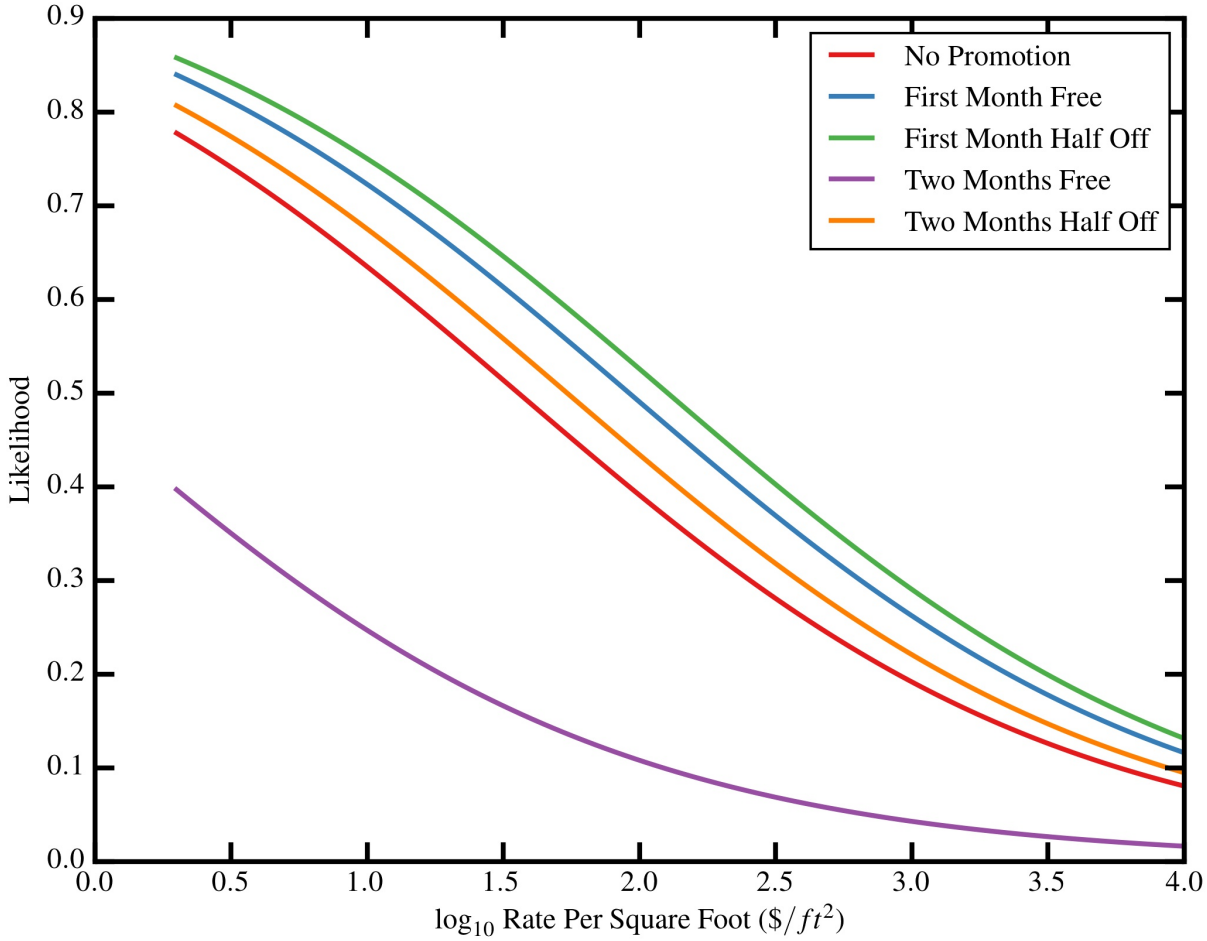

Figure 3. Likelihoods of a customer renting a unit after making a reservation as a function of the cost per square foot.

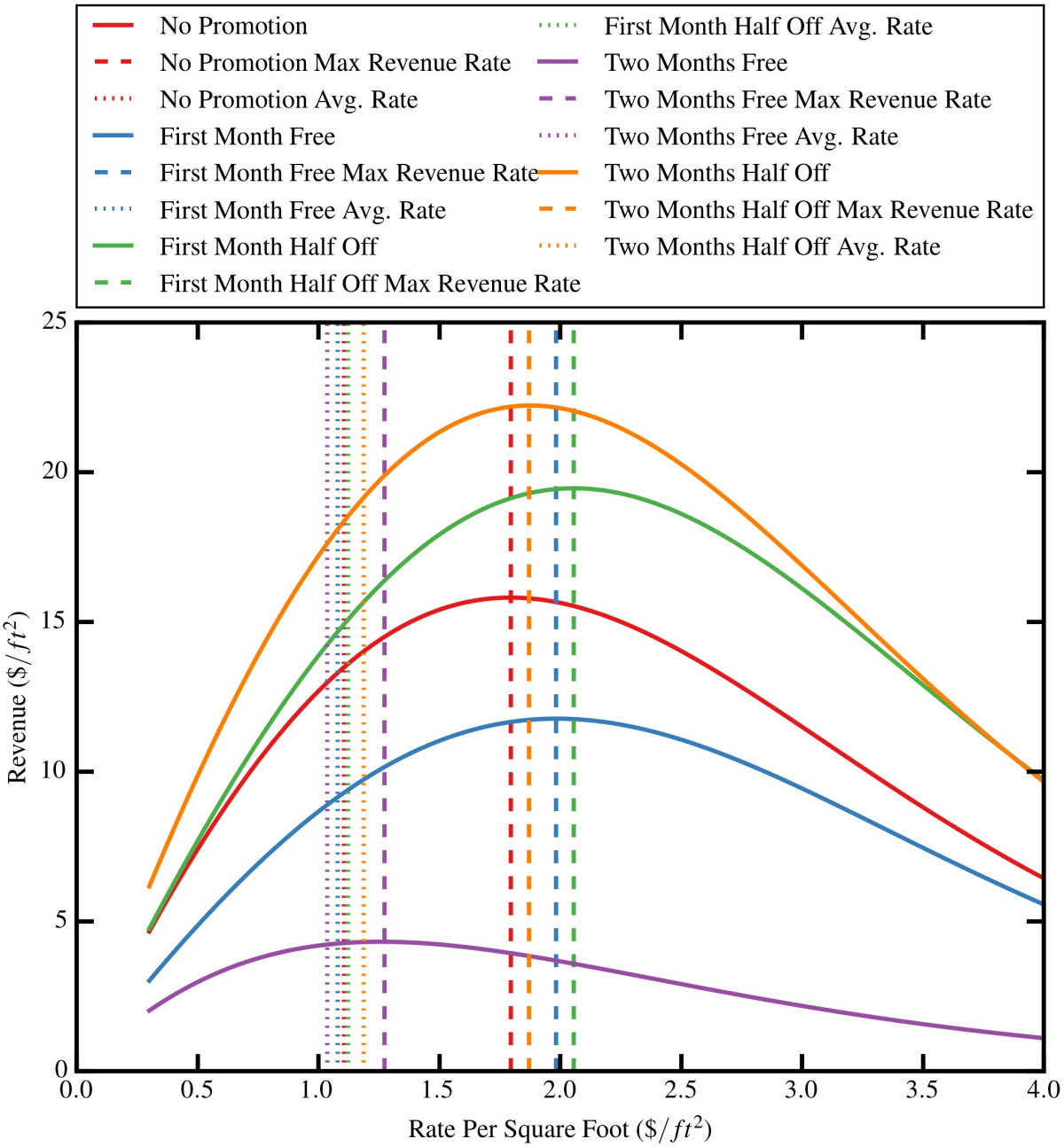

Figure 4. Revenue R generated by rentals as function of the rental rate per square foot ρ.