## **Derivation of PT Model**

## **Prospect Theory**

Prospect Theory (**PT**) is a descriptive theory of human behavior that can predict different behaviors for different individuals depending on their subjective, perceived probabilities of the outcomes of various actions and the subjective, perceived value of each outcome. The objective values of probability and value are weighted by three free parameters of the model, which are shown in equation (4) and described below. In PT, the relevant feature of a future sales price influencing whether energy units should be sold today is not its absolute value but whether it is greater or less than the current price on offer. The prospective value of a sale on day d is determined by the probability that a sale would be made at a larger or smaller price on a subsequent day. If a sale is not made at today’s price: id, the likelihood of a subsequent sale on day d+v at a smaller value implies that a sale on day d for id would be a gain and the likelihood of a subsequent sale on day d+v at a greater value implies that a sale on day d for id would be a loss. For a sales period of two days:

i-1 I

E[V2,i(n)] = ∑ w(pj)\*[n\*(i – j)]α – ∑ w(pj)\*λ[n\*(j – i)]β (4)

j=1 j=i+1

because the value of a sale at price i2 on the penultimate day, d = 2, is determined by the probability of a higher or lower price i1 being offered on the ultimate day d = 1. The first term of equation (4) computes the gain from the sale at price i2 if j = i1 < i = i2 and the second term of equation (4) computes the loss from the sale if i = i2 < j = i1. In equation (4), i = i2. j is a variable that is set to all possible values of i1.

Notice that α determines the subjective value of a gain, β determines the subjective value of a loss, and λ determines the degree of aversion to a loss. The ranges of the values of the variables are: α ≤ β ≤ 1 and λ ≥ 1. The calculation of expected value makes use of subjective probabilities w(pj) that are related to the objective probabilities used by EUT, pj so that when pj < .5, pj < w(pj) and when .5 < pj, w(pj) < pj. That is, subjective probabilities over-value lower probabilities and under-value high probabilities. Subjective probabilities are typically modeled by weighting the objective probabilities as described by Prelec, deﬁned as:

w(p**)** = exp(−(−lnp)γ), (5)

where γ = 1 reﬂects objective probability, i.e., EUT behavior. When γ < 1, the subjective probability, w(p), in (7) over-weights low probabilities and underweights high probabilities.

Notice that i and j reference the price for a single unit. If a person has n units to sell, it is assumed that the price is n \* i (no volume discounts!). The number of units influences the decision process because the effect of α < β increases with the size of the sales. That is, the perceived loss increases faster than the perceived gain. Consequently, a person might be willing to sell one unit at price i but be unwilling to sell 10 units at price i under otherwise identical conditions.  
  
In order to use PT to predict whether a person will sell or not sell when offered price i2 per unit, there must be complete information about the person’s options. That means, for all possible prices that might be offered the next day, we need to know the probability that each price might be offered. Then, as show in equation (4), we must sum over the expected values of the gains for all possible future prices lower than i and sum over the expected losses for all possible future prices lower then i.

**Fitting PT to behavior 1: Incremental change in the value of each parameter**

Now, suppose that a participant is presented with a scenario including the probabilities of all possible prices being offered the next day (so the total must equal 1) and makes a decision to sell 5 of their units at price i. To determine whether PT predicts this behavior, we must find values of α, β, λ, γ that predict a sale at 5 but not 6 units at price i. The brute force approach would be to write a program that generated all possible sets values of α ≤ β ≤ 1 and λ. The size of the space depends on the size of the incremental change in each parameter value. If we only need to consider 10 values for each parameter then the space is 104, which does not take a long time to search. However, if need to consider 1000 values for each parameter then the space is 10004, which does take a long time to search.

The size of the incremental change of a parameter is bounded by the smallest change that changes the prediction and by the largest change that changes the prediction by only one unit. For example, if an incremental change of .002 alters the prediction from a sale a maximum of 5 units to a sale at a maximum of 6 units but an incremental change of .001 does not by itself alter any prediction then .002 is the lower bound. Furthermore, if an incremental change of .003 alters the prediction from a sale a maximum of 5 units to a sale at a maximum of 7 units then .002 is also the upper bound.

There is no guarantee that a unique set of parameter values makes each possible prediction. Therefore, any search program must be written to save all sets of the parameter values that make the same prediction.

**Applying PT to a sales period to predict the probability of a sale each day**

In our experiment, participants knew the probability that each possible price might appear on each day of the sales period. From that information, they had to infer the probability that each price would appear at least once over a sales period from 69 days to 1 day. Equation (6) is the generalization of PT that performs this task.

Recall from last time, to compute the probability of a unit being sold over a multi-day sales period, we must compute the price at which it would be sold on each day of the sales period. The cut off price is the lowest price at which a unit would be sold on a given day of the sales period. As the days go by and the sales period shrinks, the cut off price must systematically decline. That is why it must be computed separately for each day even though it does not change each day. On the last day of sales period, the cut off price must be whatever is the lowest possible price that would be offered. Also, recall for a given set of parameter values, α, β, λ, and γ, one can compute the cut off price for each day of the sales period beginning on day d = 1, and moving backward to the first day, d = 68. This step must always be performed before generating the predictions, beginning with day d = 68 and moving forward to day d = 1.

## Equation (6) generalizes equation (4) to any day in the sales period:

i(d)-1 1 f+1 i’(h)-1

E[V2,i(n)] = Σ w(pj + ∑ ( ∏ Σ pk) \* pj) \* [n\*(i – j)]α –

j=i’(d-1) f=d-2 h =d-1 k=1

I 1 f+1 i’(h)-1

Σ w(pj + ∑ ( ∏ Σ pk) \* pj) \* λ[n\*(j – i)]β (6)

j=max[i’(d-1),i(d)+1] f=d-2 h =d-1 k=1

Notice that the top part of equation (6) computes the expected value of a future gain when price id is offered on day d. Therefore, the top part of equation (6) computes the probability of a sale at each price lower than id, then the expected value of the sale, and finally sums the expected values of all possible sales at a lower price on all days of the sales period.

To begin prices are considered over the range from the cut off price on day d – 1 (the next day after day) to price id – 1, the largest price less than id. Notice that lowest price in the range is the cut off price because, by definition, a sale will not be made at a price lower than the cut off price. pj is the subjective probability the price j would be offered on day d – 1, which is the probability that it would be offered on any day.

A second factor must be taken into consideration when we compute the probability that price j would be offered for sale on day d – 2. This is the probability that the units would still be available for sale on day d – 2, which is the probability that they were not sold on day d – 1. This probability is computed by the term contained by the second sigma for each of the remaining days of the sales period. From day d – 2 to day 1, the number computed under the second and third sigmas is the probability that the units have not been sold and this value is multiplied by pj to compute the probability that it was sold on that day. The cumulative probability that a sale at price j will be made on some day of the sales period is computed by the second and third sigmas.

Finally, the subjective of value of the cumulative probability that a sale at price j will be made on some day of the sales period is computed.

The same operations are performed by the bottom of equation (6) for losses.

**Fitting PT to behavior 2: Range of possible values of each parameter**

To evaluate a model, we compare the predictions of the model to a set of observations to find the set of parameter values that produces the smallest deviation between the predictions and the observations. Typically, in model development, the fit of two or more models to the observations are compared and the model whose predictions produce the smallest deviation is selected.

The number of computations necessary to find the best fit for equation (6) is more than an order of magnitude greater than the number of computations necessary to find the best fit for equation (4). First, equation (4) is implicitly designed to fit a single data point and equation (6) is explicitly designed to fit any number of data points. In our study, the number of data points to fit is 68. Second, equation (6) requires that for each set of parameter values the cut off value for a sale must be computed for all of the days in the remaining sales period. Then, for each of the 68 data points, the probability of a sale at each possible price must be computed for all of the days remaining in the sales period. Consequently, to fit the predictions generated by **one** set of parameter values to the data, on the order (682 + 68)/2 computations must be made. Therefore, if the number of sets of parameter values that must be tested is also large, the search space is very large. One way to limit the number of computations is to restrict the sets of parameter values tested to those that could possibly produce a better fit.

To begin, we assign values to our parameters: α, β, λ, and γ and predict the number of units sold for all 68 days of the sales period. This gives us six kinds of values we use to guide the search for the best fitting parameter values:

The smallest deviation so far produced between predictions and observed for all the sets of parameter values so far tested (**minDev**).

All of the sets of values of α, β, λ, and γ that have so far produced minDev.

For the most recently tested set of parameter values, for the subset of prediction, observation pairs for which prediction > observation, the sum of the differences prediction – observation (**curDev+**).

For the most recently tested set of parameter values, for the subset of prediction, observation pairs for which prediction < observation, the sum of the differences observation – prediction (**curDev-**).

**curDev** = curDev+ **+** curDev-

Notice that when only one set of parameter values has been tested, minDev = curDev.

In deciding on the next set of parameter values, we consider whether it will decrease curDev+ or decrease curDev-. We know this in advance because we know what affect that incremental increase or decrease of each parameter has on the cut off price from the structure of equation (6).

Increasing the value of α decreases the cut off price and decreasing the value of α increases the cut off price.

Increasing the value of β, λ, and γ increases the cut off price and decreasing the value of β, λ, and γ decreases the cut off price.

After we have set α, β, λ, and γ to their initial values, we vary the value of α to find the value of α that produces minDev for sets of values that also include the initial values of β, λ, and γ. Next, beginning with the values of α, β, λ, and γ that produced minDev, we vary the values of α and β to find the values of α and β that produce minDev for sets of values that also includes the initial values of λ, and γ. Next, beginning with the values of α, β, λ, and γ that produced minDev, we vary the values of α, β and λ to find the values of α, β and λ that produce minDev for sets of values that also include the initial value of γ. Finally, beginning with the values of α, β, λ, and γ that produced minDev, we vary the values of α, β, λ, and γ to find the values that produce minDev for all possible sets of values of α, β, λ, and γ.

Every time a new set of parameter values is constructed, only the value of one of the parameters of the previous set is changed. Let x be the parameter that is under consideration for an increment or decrement.

Suppose that curDev+ < curDev – minDev. Then even if the cut off price were raised sufficiently so that curDev+ = 0, the new deviation would still be greater than minDev. Therefore, then parameter x would not be changed in a direction that would raise the cut off price. Otherwise, if curDev+ ≥ curDev – minDev, then parameter x would be incremented or decremented by a sufficient amount to reduce curDev+ so that curDev+ = minDev if the value of curDev- remained unchanged.

Similarly, suppose that curDev- < curDev – minDev. Then even if the cut off price were lowered sufficiently so that curDev- = 0, the new deviation would still be greater than minDev. Therefore, then parameter x would not be changed in a direction that would lower the cut off price. Otherwise, if curDev- ≥ curDev – minDev, then parameter x would be incremented or decremented by a sufficient amount to reduce curDev- so that curDev- = minDev if the value of curDev+ remained unchanged.

Let x(1) = α, x(2) = β, x(3) = λ, and x(4) = γ. Suppose that x(i) has just been incremented or decremented and curDev has been computed. If now, neither incrementing nor decrementing x(i) can produce minDev, then whether incrementing or decrementing x(i+1) can produce minDev is considered. When x(i+1) = initial value then there is only a single value of x(i+1). However, subsequently, when x(i+1) has been set to two or more values then there are two values of x(i+1), whether to further increase the largest value of x(i+1) or further decrease the smallest value of x(i+1).

Otherwise, if i > 1, whether incrementing or decrementing x(i-1) can produce minDev is considered. Otherwise, if i = 1, whether incrementing or decrementing x(1) can produce minDev is considered.

**Fitting PT to behavior 3: Initial parameter values for each participant**

We may order the 57 participants in order of the number of days in which they sold units. Or, more precisely, we may compute the deviation between the number of units sold across the 68-day sales period for pairs of participants so that participant 1 is again the participant who sold on the fewest number of days and then order the participants so that the participant assigned to position k in the order is the participant among the 57 – (k – 1) participants not yet assigned to the ordering such that the minimum deviation between the number of units sold each day between participant k and the k – 1 participants already ordered is less than the minimum deviation for any of the other participants not yet assigned to the order. The initial parameter values for participant 1 are the ones that make PT equivalent to EUT. Then, the initial parameter values for participant k = 2 to 57 are the parameter values that produced minDev for the participant among 1 through k – 1 whose sales had the minimum deviation with the sales of participant k.

**Fitting PT to behavior 4: Finding the most efficient increment or decrement**

Suppose that one needs to increase the value of alpha, to decrease the cut off prices, to increase the number of units sold by u units. Unfortunately, I do not know an efficient procedure for finding this value. The best that I can recommend is binary search. The maximum possible value of alpha = 1. Therefore, if the current value of alpha = .k, then we compute v1 = (1 + .k) / 2, our lower range becomes .k to (1 + .k) / 2 and our upper range is (1 + .k) / 2 to 1. Suppose that after one iteration of binary search our new value of alpha is v1, the bound on our lower range is vL and the bound on our upper range is vU: vL < v1 < vU. If the number of additional units is less than u, then v2 = (v1 + vU) / 2 and the new lower and upper bounds are vL = v1 and vU. If the number of additional units is greater than u, then v2 = (v1 + vL) / 2 and the new lower and upper bounds are vL and vU = v1.

**Fitting PT to behavior 5: General considerations that should guide the search of the parameter space**

When minDev+ ~= minDev- it is unlikely that any change to the value α, β, or λ, will reduce minDev because any change that reduces minDev+ will increase minDev- and vice versa. However, if (a) minDev+ < minDev- then one may be able to improve the fit by creating a reduction in minDev- greater than the increase in minDev+. (b) If minDev+ > minDev- then one may be able to improve the fit by creating a reduction in minDev+ greater than the increase in minDev-.

(c) Even if minDev+ ~= minDev-, if the 68-day sales period can be divided into two sub-periods such that during the early part minDev+ ≤ minDev- but during the later part minDev+ ≥ minDev- then the fit of predictions to observations may be improved by increasing γ.

If (a), (b), or (c) but the fit of predictions to observations cannot be improved then the model is probably not the correct causal model because it is producing a systematic difference between predicted and observed. In the purely mathematical sense, this systematic difference can be eliminated by adding a parameter to the model. However, the new model is not a causal model.

## **Derivation of PT-TW Model**

For a time window of t days, the prospect of a gain is determined by the sum of the probabilities of a sale at a lower price than id and the prospect of a loss is determined by the sum of the probabilities of a sale at a higher price than id during the time window as shown in equation (7). Therefore, when d ≥ t:

## 

i(d)-1 t-2 i’(t-1)-1 I t-2 i’(t-1)-1

E[Vt,i(n)]=∑w(pj + ∑ (∑ pk)h \* pj)[n(i*d* – j)]α – ∑ w(pj + ∑ (∑ pi)h \* pj)λ[n(j – i*d*)]β (7)

j=i’(t-1) h=1 k=1 j=max[i’(t-1),i(d)+1] h=1 k=1

## When t > d ≥ 1: equation 6 is used to compute PT-TW.

## Notice that equation (7) is a simplified version of equation (6) in which a fixed time window replaces a shrinking sales period. When the sales period is less than the time window, PT-TW reverts to PT and equation (6) is used instead of equation (7).

## **Fitting PT-TW to behavior:**

## Equation (7) makes use of the discrete parameter x(5) = t = 68. Otherwise, the fitting procedure is the same as for PT.

## **Derivation of PT-DD and PT-DD-TW models:**

## The PT-TW models under-predicted the number of days on which a participant sold some but not all of the number of units they had available for sale. One possible reason was that PT-TW slightly over-valued gains so it predicted the sale of all units instead of just some. In fact, it is well known that the value of delayed gain is discounted in proportion to the length of the delay before the gain is made. This is called **delay discounting** (**DD**). Therefore, if participants engaged in delay discounting, it might account for why the number of some but not all sales were under-predicted. Equations 8A and 8B extend the effect of delayed discounting to the PT and PT-TW models.

## 

When d ≥ t:

i(d)-1 t-2 I

E[V2,i(n)] = Σ w(pj + ∑ (1 – ∑ ph) k \* pj) \* [n\*(i – j) / (1 + ε(k – 1)]α –

j=i’(d-1) k=1 h+j+1

I t-2 I

Σ w(pj + ∑ (1 – ∑ ph) k \* pj) \* λ[n\*(i – j) / (1 + ε(k – 1)]β (8A)

j= max[i’(d-1),i(d)+1] k=1 h+j+1

## When t > d:

i(d)-1 1 f+1 i’(h)-1

E[V2,i(n)] = Σ w(pj + ∑ ( ∏ Σ pk) \* pj) \* [n\*(i – j) / (1 + ε(f – 1)]α –

j=i’(d-1) f=d-2 h =d-1 k=1

I 1 f+1 i’(h)-1

Σ w(pj + ∑ ( ∏ Σ pk) \* pj) \* λ[n\*(j – i) / (1 + ε(f – 1)]β (8B)

j=max[i’(d-1),i(d)+1] f=d-2 h =d-1 k=1

## 

## Notice that in this version of the model, delay discounting is applied to both gains and losses. However, nearly all of the work done on delay discounting has only involved gains. The small amount of research on delay discounting of losses suggests that the effect of the delay may be different and even non-existent. Therefore, a more general test of the models would include two different k parameters: kg for gains and kl for losses. Also, it may be that (1 – k(d – f)s provides a better fit than (1 – k(d – f) because the delay function is not linear. This would add another free parameter, s, to the model and it is not clear that there are sufficient data to discriminate between these models. Also, the non-linearity may be induced by the α and β parameters.