Pandora's problem

Sam Wycherley*

Martin Weitzman is most famous for his work on the economics of climate change, but some of his early work also involved major contributions to the theory of sequential search. Weitzman (1979) characterizes the solution to "Pandora's problem".[1] This has become a classic sequential search problem with heterogeneity. A few applications:

- Sampling dealers for quoted prices in in an over-the-counter market (Duffie, Dworczak & Zhu, 2017);
- Searching for a house across different neighbourhoods;
- Job search where employers have heterogeneous characteristics;
- Technological research where there are several possible technological solutions to achieve a given goal.

Pandora's problem is as follows. Suppose there is a set $I = \{1, \ldots, n\}$ of closed boxes. Each box $i \in I$ contains a reward x_i, where x_i is a random variable with cdf F_i. The collection $\{x_i\}_{i=1}^n$ is independent. Each box costs $c_i > 0$ to open, and Pandora learns the value of the reward after a time lag $t_i \geq 0$. We have directed search – at each step, Pandora chooses whether to stop or search an additional box (at a search cost), and if continuing to search, chooses which box to search next. If stopping, Pandora selects a reward to consume from one of the boxes she has already searched. Pandora's payoff if consuming reward x at time t is $u(x,t) = e^{-rt}x$. We let $\beta_i = e^{-rt_i}$ with $r \geq 0$. For each subset $S \subseteq I$, define the *maximum sampled reward* by $y(S) = \max_{i \in S} x_i$. In particular, $y(\varnothing) = 0$.[2]

This is a dynamic programming problem that turns out to have a nice solution, as follows. Each box $i \in I$ is characterized by the triple (c_i, t_i, F_i). However, a sufficient statistic for the value of searching the box is its *reservation price* p_i. Suppose Pandora could sell the right to open box i. The reservation price of box i is the price at which she would be indifferent between selling the right and opening the box. That is, the reservation price p_i solves

$$p_i = -c_i + \beta_i \left(p_i F_i(p_i) + \int_{p_i}^{\infty} x_i F_i(\mathrm{d}x_i) \right).$$

*Email: wycherley@stanford.edu.

[1] In Greek mythology, after Prometheus stole fire and gave it to the humans, an angry Zeus created Pandora, and she brings Epimetheus (Prometheus' brother) a jar filled with all manner of plagues, as a "gift" from Zeus. When the jar was opened, the plagues were released on the humans.

[2] This is without loss – we can always normalize any outside option payoff to zero.

That is, box i's reservation price must equal the net benefit of searching box i. Rearranging, p_i solves

$$c_i = \beta_i \int_{p_i}^{\infty} (x_i - p_i) F_i(\mathrm{d}x_i) - (1 - \beta_i) p_i. \tag{1}$$

The optimal policy then reduces to a threshold policy that terminates search as soon as the maximum value Pandora has already secured exceeds her reservation price for all the remaining boxes:

Theorem 1 (Pandora's rule)**.** *In Pandora's problem, the optimal policy is characterized by the following rule:*

(a) *The next box to be opened should be the closed box with the highest reservation price, where the reservation price of box i is given by (1).*[3]

(b) *Terminate search whenever the maximum sampled reward exceeds the reservation price of every remaining box.*

Proof. First, for each box i, define

$$H_i(p_i) = \beta_i \int_{p_i}^{\infty} (x_i - p_i) F_i(\mathrm{d}x_i) - (1 - \beta_i) p_i.$$

Note H_i is continuous and strictly monotonically decreasing. Moreover,

$$\lim_{p \to -\infty} H_i(p) = \infty,$$
$$\lim_{p \to \infty} H_i(p) = \begin{cases} -\infty & \text{if } \beta_i < 1, \\ 0 & \text{if } \beta_i = 1. \end{cases}$$

Hence $\operatorname{range}(H_i) \supseteq (0, \infty)$, and since $c_i > 0$, we thus have that (1) has a unique solution $p_i \in \mathbb{R}$.

Now, suppose the policy given by (a)-(b) is optimal when there are precisely m closed boxes remaining. If $m = 1$, then immediately from the definition of the reservation value, we have that following (a)-(b) is optimal. Hence suppose there are $m + 1$ closed boxes remaining, and let S be the set of the $n - m - 1$ opened boxes. Let $j \in \arg\max_{k \in S^c} p_k$, i.e. j is a remaining unopened box in S^c with the maximal reservation value. If $y(S) \geq p_j$, then for any remaining box $i \in S^c$, the expected net gain for opening box i is $p_i - y(S) \leq p_j - y(S) \leq 0$, and so it is optimal not to open precisely one more box. Since (a)-(b) is optimal if there are m boxes remaining, it follows that not opening any more boxes is optimal if $y(S) \geq p_j$.

Conversely, if $y(S) < p_j$, then opening box j yields an expected net gain $p_j - y(S) > 0$, and so opening at least one more box is a strict improvement over stopping. Towards contradiction, suppose it is optimal to open box i first for some box i with $p_i < p_j$.

[3] In the event that more than one closed box is tied for the highest reservation value, the tie can be broken in an arbitrary way.

By assumption, following (a)-(b) is optimal after box i is opened. Let the the present expected value of opening box i and then following (a)-(b) thereafter be denoted by b.

Alternatively, let $k \in \arg\max_{h \in S^c - \{j\}} p_h$ and consider the policy of (i) opening box j; (ii) stopping if $x_j \geq p_k$; (iii) opening box i otherwise and continuing with (a)-(b) thereafter. Denote the present expected value of pursuing this policy by a.

It remains to establish $a > b$. Let

$$\begin{aligned}
\pi_j &= 1 - F_j(p_j) \\
\pi_i &= 1 - F_i(p_j), \\
w_j &= \mathbb{E}\left[x_j \mid x_j \geq p_j\right], \\
w_i &= \mathbb{E}\left[x_i \mid x_i \geq p_j\right], \\
\lambda_j &= F_j(p_j) - F_j(p_k), \\
\lambda_i &= F_i(p_j) - F_i(p_k), \\
v_j &= \mathbb{E}\left[x_j \mid p_k \leq x_j < p_j\right], \\
v_i &= \mathbb{E}\left[x_i \mid p_k \leq x_i < p_j\right], \\
\tilde{v}_j &= \mathbb{E}\left[\max\{x_j, y(S)\} \mid p_k \leq x_j < p_j\right], \\
\tilde{v}_i &= \mathbb{E}\left[\max\{x_i, y(S)\} \mid p_k \leq x_i < p_j\right], \\
\mu_i &= F(p_k) - F(p_i), \\
u_i &= \mathbb{E}\left[x_i \mid p_i \leq x_i < p_k\right], \\
d &= \mathbb{E}\left[\max\{x_i, x_j, y(S)\} \mid p_k \leq x_i < p_j;\ p_k \leq x_j < p_j\right], \\
\Phi &= \mathbb{E}\left[\Psi(S^c - \{i,j\}, \max\{x_i, x_j, y(S)\}) \mid x_i < p_k;\ x_j < p_k\right],
\end{aligned}$$

where $\Psi(E, \bar{y})$ is defined to be the state value function if the current maximum sampled reward is $\bar{y}$ and the set of closed boxes remaining is E.[4]

Now, if following the policy involving opening box i first, the present expected value is

$$\begin{aligned}
b = -c_i &+ \beta_i \pi_i w_i + \beta_i \lambda_i [-c_j + \beta_j \pi_j w_j + \beta_j \lambda_j d + \beta_j (1 - \pi_j - \lambda_j) \tilde{v}_i] \\
&+ \beta_i (1 - \pi_i - \lambda_i)[-c_j + \beta_j \pi_j w_j + \beta_j \lambda_j \tilde{v}_j] + (1 - \pi_i - \lambda_i)(1 - \pi_j - \lambda_j) \beta_i \beta_j \Phi.
\end{aligned}$$

Under the alternative policy (opening box j before i), the present expected value is

$$\begin{aligned}
a = -c_j &+ \beta_j \pi_j w_j + \beta_j \lambda_j \tilde{v}_j + \beta_j (1 - \pi_j - \lambda_j)[-c_i + \beta_i \pi_i w_i + \beta_i \lambda_i \tilde{v}_i] \\
&+ (1 - \pi_i - \lambda_i)(1 - \pi_j - \lambda_j) \beta_i \beta_j \Phi.
\end{aligned}$$

[4] Recursively, Ψ is given by $\Psi(\varnothing, \bar{y}) = \bar{y}$ and

$$\Psi(E, y) = \max\left\{\bar{y}, \max_{i \in E}\left\{-c_i + \beta_i \left(\Psi(E - \{i\}, \bar{y}) F_i(\bar{y}) + \int_{\bar{y}}^{\infty} \Psi(E - \{i\}, x_i) F_i(\mathrm{d}x_i)\right)\right\}\right\}$$

for $E \neq \varnothing$.

Combining, and using (1) to substitute for c_i and c_j, we have

$$\begin{aligned}
a - b &= (c_j - \beta_j\pi_j w_j)(\beta_i(1-\pi_i) - 1) + (c_i - \beta_i\pi_i w_i)((1 - \beta_j(1-\pi_j-\lambda_j))) \\
&\qquad + \beta_j\lambda_j\tilde{v}_j - \beta_i\beta_j\lambda_i\lambda_j d - \beta_i\beta_j(1-\pi_i-\lambda_i)\lambda_j\tilde{v}_j \\
&= (p_j - p_i)(1-\beta_j(1-\pi_j))(1-\beta_i(1-\pi_i)) \\
&\qquad + (v_i - p_i)\beta_i\lambda_i(1-\beta_j(1-\pi_j)) + (\tilde{v}_j - z_i)\beta_j\lambda_j(1-\beta_i(1-\pi_i)) \\
&\qquad + (u_i - p_i)\beta_i\mu_i(1-\beta_j(1-\pi_j-\lambda_j)) \\
&\qquad + (\tilde{v}_j + v_i - p_i - d)\beta_i\beta_j\lambda_i\lambda_j. \qquad (2)
\end{aligned}$$

Now by definition of d,

$$\begin{aligned}
d &= p_k + \mathbb{E}\left[\max\{\max\{x_j, y(S)\} - p_k, x_i - p_k\} \mid p_k \le x_i < p_j;\, p_k \le x_j < p_j\right] \\
&\le p_k + \mathbb{E}\left[\max\{x_j, y(S)\} - p_k + x_i - p_k \mid p_k \le x_i < p_j;\, p_k \le x_j < p_j\right] \\
&= \tilde{v}_j + v_i - p_k \\
&\le \tilde{v}_j + v_i - p_i.
\end{aligned}$$

From this inequality and $\beta_i, \beta_j \in (0,1]$, we have that every term of (2) is nonnegative. Moreover, the first term of (2) is strictly positive. Thus $a > b$. It follows that sampling remaining boxes in a manner not consistent with (a) cannot be optimal.

Thus any optimal policy must necessarily take the form (a)-(b). Since an optimal policy exists by nature of the finite dynamic programming problem, (a)-(b) is also sufficient to characterize optimal policy. □