Stop Waiting Problem: Decision Rule with $\Psi$ function and Application with Share Prices

Suggested Citation: Kohn, Wolfgang (2014) : Stop Waiting Problem: Decision Rule with $\Psi$ function and Application with Share Prices, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:
http://hdl.handle.net/10419/93096
Stop Waiting Problem
Decision Rule with $\Psi$ function and Application with Share Prices

Wolfgang Kohn*
University of Applied Sciences, Bielefeld
March 4, 2014

Abstract
In this paper the stop-waiting strategy of Franz Bruss [1] is set into a simple probabilistic framework and applied to the apple share prices from 1984 to 2013. Within the probabilistic framework a heuristic and a mathematical decision rule using the $\Psi$ function is developed. The results are in line with Bruss’s theory. We apply the stop-waiting strategy to the Apple share prices and compare the results with a simple start-end and chart technique strategy.

1 When Should One Stop Waiting?

This kind of question appears in real cases, when one wants to sell (or buy) something. How many offers (opportunities) should one let pass by before a decision should be taken

* wolfgang.kohn@fh-bielefeld.de
JEL Classification: C65, E47, G17; AMS Classification: 60G40, 62L15
for the next offer which is superior? We assume $n$ independent events. Before we take a decision, we wait for $j$ events and take the next event which is better than all before.

**Example 1.1**

In the following small simulation ten cases ($w = 10$) are simulated each with $n = 5$ events. In the first case ($w = 1$) one has to stop after the first or second event to get the best $x = 5$. Stop waiting after the first event will lead to $x = 5$, because $x = 1$ is not superior against $x = 2$. Stop waiting after the second event lead directly to the best $x = 5$. Waiting longer will lead to the situation that no decision is taken, because nothing is better than $x = 5$. In case $w = 5$ one has to wait until $j = 2$ to receive the best event.

Which $j$ is connected with the highest probability for the best event? Counting out the cases on which $j$ the best event occurs shows, that this turns out for $j = 2$ in this small simulation. So stop waiting at $j = 2$ is connected to the highest relative frequency. What is the theory behind?

<table>
<thead>
<tr>
<th>$w$ = 1</th>
<th>$w$ = 2</th>
<th>$w$ = 3</th>
<th>$w$ = 4</th>
<th>$w$ = 5</th>
<th>$w$ = 6</th>
<th>$w$ = 7</th>
<th>$w$ = 8</th>
<th>$w$ = 9</th>
<th>$w$ = 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>$j$ = 1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>$j$ = 2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>$j$ = 3</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>$j$ = 4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>$j$ = 5</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>

0.5 0.6 0.4 0.3

We define $A_k$ as the best event. $k$ is the index for the position of the best event.

\[
\begin{align*}
&j \\
&1 \longrightarrow \ldots \longrightarrow k-1 \longrightarrow k \longrightarrow \ldots \longrightarrow n \\
&B_j &\quad A_k
\end{align*}
\]

**Figure 1: Stop Waiting**

$B_j$ is defined as an inferior event in comparison to $A_k$ in $j = 1 \ldots k - 1$, but best event in $j$. We call this kind of event some times for simplicity the second best event. It is not necessarily a second best event. For example if 5 is the best event and we have a series 2, 3, 5, so 3 is inferior to 5, but the „second best event“ in $j = 2$. $j$ is the index for the events during the waiting phase.
2 Conditional Probability for the Best Event

The probability for the best event $A_k$ AND an inferior event $B_j$ in $j$ is denoted with $\Pr(A \cap B_j)$. We develop this probability within four steps.

1. We start looking on the conditional probability

$$Pr(A_k \mid B_j) = \frac{Pr(A_k \cap B_j)}{Pr(B_j)}$$

This is the probability for the best event $A_k$ on position $k$ if $B_j$ the second best event occurs during the waiting time. To receive $A_k$ as the best event $k - j - 1$ events must be inferior to $A_k$. In $j$ (waiting period) $k - 1$ events appear. The conditional probability not to receive $A_k$ in $j$ is therefore

$$Pr(\overline{A}_k \mid B_j) = \frac{k - j - 1}{k - 1}$$

Consequently the Laplace probability for $A_k$ given $B_j$ is $1 - Pr(\overline{A}_k \mid B_j)$.

$$Pr(A_k \mid B_j) = \begin{cases} \frac{j}{k - 1} & \text{for } k = 2, \ldots, k - 1 \text{ and } j < k \\ 0 & \text{for } j \geq k \end{cases}$$

Example 2.1

$n = 5$. In the following table the probabilities for $Pr(A_k \mid B_j)$ are calculated.

<table>
<thead>
<tr>
<th>$j$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k$</td>
<td>NA</td>
<td>NA</td>
<td>0.5</td>
<td>0.3333333</td>
<td>0.25</td>
</tr>
<tr>
<td>$k$</td>
<td>2</td>
<td>NA</td>
<td>1.0</td>
<td>0.6666667</td>
<td>0.50</td>
</tr>
<tr>
<td>$k$</td>
<td>3</td>
<td>NA</td>
<td>NA</td>
<td>1.0000000</td>
<td>0.75</td>
</tr>
<tr>
<td>$k$</td>
<td>4</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>1.00</td>
</tr>
</tbody>
</table>

2. The probability for an inferior event $Pr(B_j)$ in $j$. The chance to receive the inferior event in $n$ events is

$$Pr(B_j) = \frac{1}{n} \text{ for } j = 1, \ldots, n - 1$$

Example 2.2 (cont.)

Probability of $Pr(B_j)$.

<table>
<thead>
<tr>
<th>$j$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>j= 1</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
</tbody>
</table>
3. Now we are in the position to calculate the probability \( \Pr(A_k \cap B_j) \). We rewrite the equation (1).

\[
\Pr(A_k \cap B_j) = \Pr(A_k \mid B_j) \Pr(B_j) = \frac{j}{k-1} \frac{1}{n} \quad \text{for} \ j < k
\]

**Example 2.3 (cont.)**

*In our example the probabilities for \( \Pr(A_k \cap B_j) \) are*

<table>
<thead>
<tr>
<th>( j )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( k )</td>
<td>NA</td>
<td>0.2</td>
<td>0.1</td>
<td>0.06666667</td>
<td>0.05</td>
</tr>
<tr>
<td>( k )</td>
<td>NA</td>
<td>0.2</td>
<td>0.13333333</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>( k )</td>
<td>NA</td>
<td>NA</td>
<td>0.20000000</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>( k )</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>0.20</td>
<td></td>
</tr>
</tbody>
</table>

4. The probability of \( \Pr(A \cap B_j) \), which is the summation over \( k = j + 1, \ldots, n \) of all probabilities \( \Pr(A_k \cap B_j) \) for a specific \( j \) is consequently

\[
\Pr(A \cap B_j) = \sum_{k=j+1}^{n} \Pr(A_k \cap B_j) = \frac{1}{n} \sum_{k=j+1}^{n} \frac{j}{k-1} \quad \text{for} \ j = 1, \ldots, n - 1
\]

**Example 2.4 (cont.)**

*Probabilities of \( \Pr(A \cap B_j) \).*

<table>
<thead>
<tr>
<th>( j )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.41666667</td>
<td>0.43333333</td>
<td>0.35000000</td>
<td>0.20000000</td>
<td></td>
</tr>
</tbody>
</table>

We see that with \( n = 5 \) the best strategy is to wait for \( j = 2 \) opportunities and then to take the next one which is superior. This will lead to the best event with a probability of \( \Pr(A \cap B_2) = 0.433 \).

### 3 Decision Rule for Stop Waiting with \( \Psi \) function

We search for a decision rule for \( j \) where \( \Pr(A \cap B_j) \) has a maximum within a finite range. A heuristic approach is to wait as long as the probability for \( j \) inferior events is smaller than \( \Pr(A \cap B_j) \)

\[
j \Pr(B_j) \leq \Pr(A \cap B_j) \quad (3)
\]

The last time this condition holds is the best position for stop waiting: \( j_{-\text{sup}} \). Rewriting this condition leads to

\[
\sup_{j=1,\ldots,n-1} \left\{ j : j \Pr(B_j) \leq \Pr(A \cap B_j) \right\} \Leftrightarrow \sup_{j=1,\ldots,n-1} \left\{ j : 1 \leq \sum_{k=j+1}^{n} \frac{1}{k-1} \right\}
\]

\[
(4)
\]
Waiting for $j$ inferior events must be less likely than obtaining the best event and an inferior event. Graphically the lower bound $j \Pr(B_j)$ is for $n$ given a straight line from $\frac{1}{n}, \frac{2}{n}, \ldots, 1 - \frac{1}{n}$ in steps of $\frac{1}{n}$. As long as the points are below the curve of $\Pr(A \cap B_j)$ it is worthwhile to wait. The last point below $\Pr(A \cap B_j)$ indicates stop waiting. This is the point of $j_{\text{sup}}$ (see fig. 2, orange line and orange point).

$$\text{eq. 2}$$ $\text{eq. 3}$ $\text{eq. 4}$ $\text{eq. 5}$ $\text{eq. 6}$ $\text{eq. 7}$ $\text{eq. 8}$

A more mathematical approach is to find the maximum via differentiation. To do this we rewrite

$$\Pr(A \cap B_j) = \frac{1}{n} \sum_{k=j+1}^{n} \frac{j}{k-1} = \frac{j}{n} \sum_{k=1}^{n-1} \frac{1}{k}$$

The second sum could now be expressed with the psi function $\Psi(n)$ (or digamma function). One recursive definition of $\Psi(n)$ is $\sum_{k=1}^{n-1} \frac{1}{k} + \Psi(1)$ with

$$\Psi(1) = \lim_{n \to \infty} \left( \sum_{k=1}^{n} \frac{1}{k} - \ln k \right) = -\gamma \approx -0.5772\ldots$$

So $\Psi(n) - \Psi(1)$ represents the sum $\sum_{k=1}^{n-1} \frac{1}{k}$ and $\Psi(n) - \Psi(j)$ with $j = 1, \ldots, n - 1$ is the series $\sum_{k=1}^{n-1} \frac{1}{k}, \sum_{k=2}^{n-1} \frac{1}{k}, \ldots$ (see [5, Chap. 12]).

$$\text{Pr}(A \cap B_j) = \frac{j \Psi(n) - j \Psi(j)}{n} \quad \text{(5)}$$

The $\Psi(n)$ function is continuous (!) and could be differentiated with respect to $j$ and set to zero to find the necessary condition for an extreme value (in fig. 2 green curves, blue connected points are $\Pr(A \cap B_j)$ evaluated at $j$).

$$\frac{\partial \Pr(A \cap B_j)}{\partial j} = \frac{\Psi(n) - \Psi(j) - j \Psi'(j)}{n} = 0 \quad \text{with} \quad \Psi'(j) = \frac{d\Psi(j)}{dj} \quad \text{(6)}$$
This derivative is a monotonically decreasing function with a unique real root \( j_{\text{sup}} \). For \( j < 4 \) there are no roots of (6). Rounding \( j_{\text{sup}} \) to the next integer leads to the position where we should stop waiting: \( \lceil j_{\text{sup}} + 0.5 \rceil \). This point indicates the maximum of \( \Pr(A \cap B_j) \).

Rewriting eq. (6)

\[
\frac{\Psi(n) - \Psi(j)}{n} \equiv \frac{j \Psi'(j)}{n}
\]

and expanding the equation with \( j \) leads to

\[
\Pr(A \cap B_j) = j \Pr(B_j) j \Psi'(j)
\]

\( j \Psi'(j) \) is a monotonically decreasing function to 1. Removing \( j \Psi'(j) \) from eq. (7) we abtain

\[
\Pr(A \cap B_j) \geq j \Pr(B_j)
\]

which is the condition (3). As long as \( \frac{\Pr(A \cap B_j)}{\Pr(B_j)} \geq j \) we should wait.

We see in fig. 2 that for increasing \( n \) the discrete function of \( \Pr(A \cap B_j) \) in eq. (2) convergences to the continuous function of \( \Pr(A \cap B_j) \) in eq. (5).

### 4 Simulation and Convergence of the Probability for the Best Event

We generate 1000 repeated samples of \( n = 40 \) and \( n = 100 \) random numbers without replacement. We clearly see in fig. 3 (left, dashed lines simulated probabilities) that the theoretical consideration from eq. (2) holds. The chance for the best event converges to \( e^{-1} \) which is simulated fig. 3, right. \( j_{\text{sup}} \) convergences to \( ne^{-1} \), so the relative position is \( e^{-1} \) (see [2]). To see this we approximate

\[
\sum_{k=j+1}^{n} \frac{1}{k-1} \approx \int_{j}^{n} \frac{1}{k} \, dk = -\ln \left( \frac{j}{n} \right)
\]

Now we can write for \( \Pr(A \cap B_j) \) (see fig. 2, green dashed line for the approximation)

\[
\Pr(A \cap B_j) = j \frac{1}{n} \sum_{k=j+1}^{n} \frac{1}{k-1} \approx - j \frac{1}{n} \ln \left( \frac{j}{n} \right)
\]

(8)

The extremum of (8) appears on

\[
\frac{\partial \Pr(A \cap B_j)}{\partial j} = - \frac{1}{n} - \frac{1}{n} \ln \left( \frac{j}{n} \right) = 0 \Rightarrow j_{\text{sup}} \approx \frac{n}{e}
\]

and \( \sup \Pr(A \cap B_j) \) for \( j_{\text{sup}} = ne^{-1} \) convergences to \( e^{-1} \) for \( n \to \infty \) (see fig. 3). So it is possible to find an approximate decision rule for a large undefined number of events.
5 Application

We take the Apple share prices from 9th December 2010 to 9th December 2013 (n = 755) and compare 3 strategies of selling:

1. stop waiting
2. selling at the end of period
3. chart technique: 38 days moving average crossing the first time from above 100 days moving average (see [6, Chap. 4])

The profit is the difference between selling price and first price. If no selling point is indicated (1st and 3rd strategy) the profit is the difference of the last price in the window to the 1st price. First we use the full period and compute the profit (see fig. 4, left). Second we move over the historical data with a window of 250 non overlapping days and compute an average profit. Third we use the log return, cut off all returns above and below 1.1% change to avoid a drift in the simulated share prices and take 100 random samples (resampling the log return) to simulate share prices with the starting price of 9th December 2010 (see fig. 4, right top). We apply the 3 strategies to the full period and the moving 250 days window. Fourth we simulate share price with a drift by cutting off only returns above and below 1.5% see fig. 4, right bottom) and compute the profits.

1. Within the full period a best-selling point is indicated at the 18th January 2012 (≈ $ = \frac{755}{2} \approx 278$). The point is connected to a higher price than all past prices (the next superior price after stop waiting), but far away from the peak price (see fig. 4, left). The profit is $109.35$ (full period) and the profit moving with a 250 days window over the dataset is on average $102.61$. The simulation with 100 samples
with no drift shows for the full period an average a loss of $−6.01 ± 42.31 (standard error) and a loss of $−2.7 ± 14.3 (standard error of mean) for the series of 250 days. The simulation with drift a profit turns out for the full period ($64.33 ± 41.13) as well as for the 250 days periods ($23.9 ± 16.11).

2. The simple start-end strategy is for the historical dataset for the full period the best: the profit is $246.67. The reason is simply the upward trend. Moving with a 250 days window over the data reduces the: $90.2. Using the simulated prices with no drift an average loss of $−8.96 ± 46.26 and for the 250 days series an average loss $−3.05 ± 15.5 appears. The simulated data with drift shows again the highest profits: full period $131.5 ± 88.62, 250 days $43.64 ± 29.96.

3. The chart strategy shows up a profit of $10.91 (full period) and on average for the moving 250 days window with $39.29. In the simulation with no drift the average profit turns as well into a loss, but surprisingly it is the smallest: full period $−0.6 ± 16.64 and 250 days $−0.17 ± 8.22. The results for the simulation with drift shows in line with the previous results the smallest profits: full period $16.94 ± 28.23 and 250 days $17.57 ± 17.28.

To sum up no strategy seems to be superior (see tab. 1). As long as a drift is present the simple start-end strategy seems to be favourite. For a 250 days trade the stop waiting shows up the best result (with one exception for the simulation with drift 250 days). And the chart technique seems to minimize losses. Of course the result depends on the historical dataset and on the random samples.
<table>
<thead>
<tr>
<th>Strategy</th>
<th>Historic Data</th>
<th>Simulated Data</th>
<th>Simulated Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full Period</td>
<td>250 Days</td>
<td>Full Period</td>
</tr>
<tr>
<td>Stop waiting</td>
<td>109.35</td>
<td>102.61</td>
<td>-6.01</td>
</tr>
<tr>
<td></td>
<td>±42.31</td>
<td>±14.3</td>
<td>±41.13</td>
</tr>
<tr>
<td>Start-End</td>
<td>246.67</td>
<td>90.2</td>
<td>-8.96</td>
</tr>
<tr>
<td></td>
<td>±46.26</td>
<td>±15.5</td>
<td>±88.62</td>
</tr>
<tr>
<td>Chart</td>
<td>10.91</td>
<td>39.29</td>
<td>-0.6</td>
</tr>
<tr>
<td></td>
<td>±16.64</td>
<td>±8.22</td>
<td>±28.23</td>
</tr>
</tbody>
</table>

Table 1: Historical and Simulated Profits Using Different Strategies

6 Conclusion

The stop strategy shows a simple rule how to find under uncertainty a moment for stop waiting. But the probability is clearly below 50 percent, often below 40 percent to achieve the best event. So the chance to fail is greater than to succeed. A probability cannot be interpreted as a number which ensures anything. Applied to our above results we know that on average the indicated position is optimal within the theory, but not in a specific case. Furthermore the above results can very much be influenced by real effects which the theory cannot take into account, but it looks like, that there is no optimal strategy regarding the application.

References


