

# Identifying consideration sets through search and an implementation using Shiny

---

Erlend Dancke Sandorf<sup>a</sup> Danny Campbell<sup>a</sup>

<sup>a</sup>Economics Division, Stirling Management School, University of Stirling

ENVECHO Verona 2019

# The search for information - the search for alternatives

---

# The role that information and search plays in choice has a long history in economics

- Search models – Simon (1955), Stiegler (1961), Weitzman (1979), Gabaix et al. (2006)
- Consideration set models – Richardson (1982), Roberts and Lattin (1991)
- Information acquisition models – Hausmann and Laje (2008), Chorus et al. (2013)

# The role that information and search plays in choice has a long history in economics

- Search models – Simon (1955), Stiegler (1961), Weitzman (1979), Gabaix et al. (2006)
- Consideration set models – Richardson (1982), Roberts and Lattin (1991)
- Information acquisition models – Hausmann and Laje (2008), Chorus et al. (2013)

People search as long as the expected gain from search exceeds the marginal cost

In many (if not most) choice situations, options are evaluated sequentially



# In many (if not most) choice situations, options are evaluated sequentially

skyscanner

Flights | Hotels | Car Hire

Edinburgh (EDI) - Kobe (UKB)

1 adult | Economy

Show whole month

Additional bag fees may apply

Get Price Alerts

Travel tip

Fly to Osaka Kansai International and save £116 per person.

Stops

Direct  
None

1 stop  
None

2+ stops  
£754

Departure times

Outbound  
00:00 - 23:59

Journey duration

31 hours - 48.5 hours

Airlines

Select all | Clear all

oneworld (none)

SkyTeam (none)

Star Alliance (none)

Value Alliance (none)

24 results sorted by Best

Sort by Best

Best	Cheapest	Fastest
£754 30h 40	£754 30h 40	£3,582 20h 35

JR JRCENTRAL

"Central Japan Shinkansen/Train Portal" DEBUT!

Mt. Fuji, Kyoto, Takayama, and more! Book shinkansen/train or find tours and rail passes!

Check prices and save

Select →

easyJet + Qatar Airways + Skymark Airlines

17:55 EDI → 08:35<sup>+2</sup> UKB

30h 40  
3 stops  
CPH, DOH, HND

1 deal  
£754

Select →

Self-transfer

easyJet + Qatar Airways + ANA (All Nippon Airways)

17:55 EDI → 07:35<sup>+2</sup> UKB

29h 40  
3 stops  
CPH, DOH, HND

1 deal  
£812

Select →

Self-transfer

Emirates + Hawaiian Airlines

20:15 → 08:35<sup>+2</sup>

28h 20  
1 stop

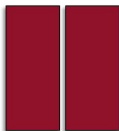
1 deal  
£666

This means that consideration sets grow sequentially  
with each period of search

t = 1



t = 2



t = 3



t = 4



# Gathering data to test hypotheses with respect to search and satisficing

---



A good that is familiar, reasonably priced and induces people to search



# Selected attributes: Country of Origin, Color, Alcohol Content, Grape Variety, Characteristic, Organic and Price



# We used a web scraping algorithm to gather information on our attributes

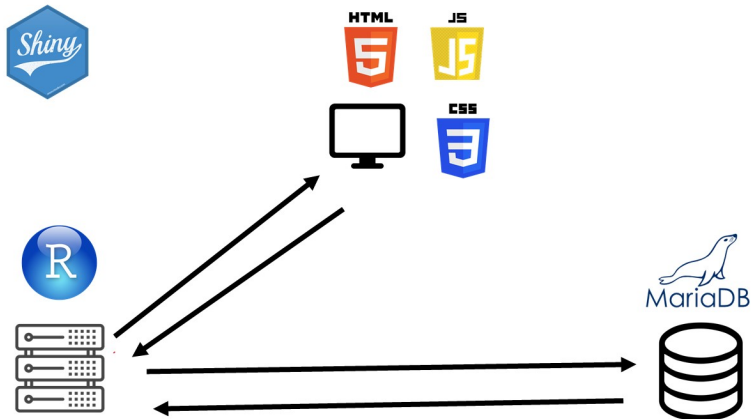
The screenshot shows a web browser with the Sainsbury's website. The search bar contains 'red wine'. The search results page displays various wine products. An RStudio console window is overlaid on the page, showing R code used for web scraping. The code includes the following steps:

```
12 # Grab the links of all the wines on the Tesco website
13 url_base <- "https://www.tesco.com/groceries/en-GB/shop/drinks/wine/all?pages"
14
15 # Define counter and progress bar.
16 counter <- 0
17 progress_bar <- utils::txtProgressBar(min = 0, max = 28, style = 3)
18
19 page_urls <- lapply(seq_len(28), function(i) {
20   # A counter for debugging and progress bar
21   counter <-> counter + 1
22   utils::setTxtProgressBar(progress_bar, counter)
23 })
24
25 # Grab the website URL
26 web_page <- read_html(paste0(url_base, i))
27 url <- web_page %>%
28   html_nodes("a.sc-htadjsa") %>%
29   html_attr("href") %>%
30   deparse1()
31   lapply(urls, function(x) {
32     x[1]
33   })
34
35 # Close the progress bar
36 close(progress_bar)
37
38 # Put page URLs in a vector
39 page_urls <- do.call(c, page_urls)
40
41 # Grab the information from the website
42 url_base <- "https://www.tesco.com"
43
44 wine_info_id <- c("region-of-origin", "wine-colour", "#alcohol-units",
45                 "#percentage-alcohol", "#country", "#alcohol-type",
46                 "#net-contents", "#grape-variety")
47
48 name_str <- c("wine", "price", "region", "color", "units",
49             "alcohol", "country", "type", "content", "grape")
50
51 # Define counter and progress bar
52 counter <- 0
53 progress_bar <- utils::txtProgressBar(min = 0, max = length(page_urls), style = 3)
54
55 data_tesco <- lapply(seq_along(page_urls), function(i) {
```

# We sample alternatives according to how common they are in the store

	A	B	C	D	E	F	G	H
1	country	color	alcohol	grape	characteristic	organic	price	weight
2	New Zealand	White	11.50%	Pinot Grigio	2 - Semi dry	No	8	3.02E-15
3	South Africa	Red	13%	Tempranillo	5 - Full bodied	Yes	4.5	2.39E-15
4	Chile	Red	14%	Malbec	5 - Full bodied	Yes	5.5	5.26E-16
5	New Zealand	Rosé	10.50%	Merlot	1 - Dry	Yes	5.5	2.62E-17
6	Italy	White	14.50%	Blend	1 - Dry	Yes	15	1.63E-16
7	South Africa	Red	11%	Malbec	5 - Full bodied	Yes	5.5	3.56E-16
8	USA	Rosé	12%	Pinot Noir	3 - Medium dry	No	5.75	2.83E-16
9	Italy	Rosé	14%	Tempranillo	4 - Semi sweet	Yes	9.5	7.08E-17
10	Australia	Red	11.50%	Blend	1 - Light bodied	Yes	7.25	1.12E-15
11	New Zealand	Rosé	13%	Cabernet Sauvignon	3 - Medium dry	No	16.5	6.16E-17
12	Italy	Red	12%	Blend	1 - Light bodied	Yes	8.5	6.13E-16
13	France	White	10%	Sauvignon Blanc	3 - Medium dry	No	5	4.96E-16
14	USA	White	12%	Pinot Grigio	2 - Semi dry	Yes	8.75	4.44E-16
15	New Zealand	Rosé	12%	Pinot Noir	4 - Semi sweet	Yes	5	3.84E-16
16	New Zealand	White	13%	Pinot Grigio	1 - Dry	Yes	17.5	2.89E-16
17	South Africa	Red	12.50%	Merlot	4 - Full bodied	No	6.5	1.07E-15
18	Italy	Rosé	14%	Merlot	3 - Medium dry	Yes	6.25	7.08E-17
19	New Zealand	Red	15%	Tempranillo	4 - Full bodied	Yes	11	2.19E-17
20	Italy	Rosé	12%	Pinot Noir	4 - Semi sweet	No	5.5	3.67E-16

# We implement the survey using Shiny and host it on shinyapps.io





Question 1: Please choose your most preferred wine

	None of these wines	Wine 1	Wine 2	Wine 3
Country		Australia	Chile	Australia
Color		White	White	Red
Alcohol		12%	13.5%	12.5%
Grape		Chenin Blanc	Blend	Tempranillo
Characteristic		1 - Dry	3 - Medium dry	5 - Full bodied
Organic		No	Yes	Yes
Price		5	9	13
I choose	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>





[www.acrg.site](http://www.acrg.site)

Economics Division

Stirling Management School

University of Stirling

Scotland

# Econometric model

---



# Utility can be described by a separable and additive utility function

$$U_{nis} = \beta X_{nis} + \varepsilon_{nis}$$

$U_{nis}$  Utility

$\beta$  Vector of parameters to be estimated

$X_{nis}$  Levels of the attributes

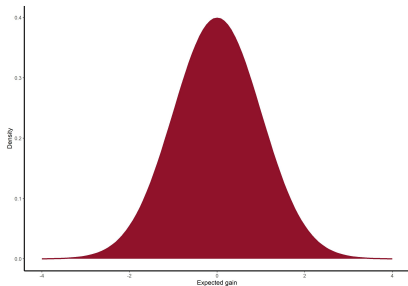
$\varepsilon$  Type I Extreme value distributed error term with variance  $\pi^2/6$

The possible gain from search is the difference between any alternative and the current best

$$g = u - u_{\max}$$

The value of all possible gains is the area under the “gain” curve

$$\bar{g} = \int_{-\infty}^{+\infty} gP(g)dg,$$



# With recall you cannot lose utility by searching for another alternative

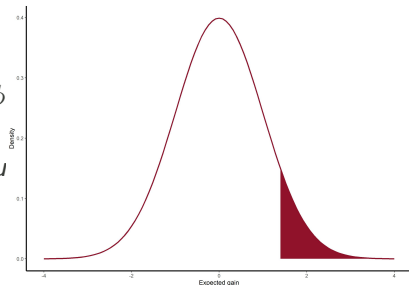
$$g = \begin{cases} u - u_{\max} & \text{if } u \geq u_{\max} \\ 0 & \text{if } u \leq u_{\max} \end{cases}$$

# The gain from searching is the area under the “gain” curve above the current best

$$\begin{aligned}\bar{g} &= \int_{u_{\max}}^{+\infty} (u - u_{\max}) \phi(u) du \\ &= \int_{u_{\max}}^{+\infty} u \phi(u) du - \int_{u_{\max}}^{+\infty} u_{\max} \phi(u) du \\ &= \phi(u_{\max}) - u_{\max} \int_{u_{\max}}^{+\infty} \phi(u) du\end{aligned}$$

where

$$u_{\max} = (U_{\max} - \mu_t) / \sigma_t$$



# An individual will search as long as expected gains are higher than the marginal cost of searching

$$\bar{G} - \bar{c} > 0$$

where

$$\bar{G} = \bar{g}\sigma$$

i.e. the non-standardized gain to be compared with the marginal cost of search  $\bar{c}$ , e.g. time, money, cognitive cost of maintaining a consideration set

$$P(i_s | C_{ns}) = \prod_{t=1}^{T=J} \left[ \frac{\exp(\beta x_{nis})}{\sum_{j \in C_{ns}^t} \exp(\beta x_{njs})} \right]^{l_t}$$

where

$$l_t = \begin{cases} 1 & \text{if } \bar{G}_t - \bar{c}_t < 0 & t = t^* \\ 0 & \text{if } \bar{G}_t - \bar{c}_t \geq 0 & \forall t \neq t^* \end{cases}$$

and  $t^*$  is the first time the condition is TRUE.