Social Proof Theory

Andrew Scutt[1]

[1]Faculty of Psychology University of Toronto

Social Proof Theory, December 2025

Table of Contents

Table of Contents

As a social species, humans often resort to efficient social learning mechanisms to solve problems. In particular, observational learning, the process of acquiring new knowledge by watching others, is quintessential to human learning (Bandura 1975).

An interesting subset of observational learning, social proof theory, posits that individuals in unfamiliar situations are more likely to mimic the behaviour of others when they perceive their entourage possesses more contextual knowledge than them (Cialdini 1984).

However, past research has not explored other determinants that may engender behaviour reflective of social proof theory, such as cognitive cost and agreement.

Table of Contents

Setup:

- You are looking to buy a laptop with a specific set of characteristics: $C_n = (y_{n,1}, x_{n,1}, x_{n,2}, x_{n,3})$ where
  - n is the computer's identification number
  - $y_{n,1}$ is computer n's price
  - $x_{n,1}$ is computer n's CPU type
  - $x_{n,2}$ is computer n's RAM specification
  - $x_{n,3}$ is computer n's SSD specification
- You have to offer a price (o) to a clerk who will either accept or deny your offer depending on whether it exceeds or meets their WTA (τ)

- You have two options:
  - **Computer Sampling**: Derive the optimal offering price by perusing the current laptop market and trying to figure out how each characteristic contributes to the overall market price
  - **Social Sampling**: You can look at what other people have offered for the same laptop you are trying to purchase

Regarding the seller's WTA, this value will be pulled from the following distribution: $\tau \sim \mathcal{N}(\mu_\tau, \sigma_{\tau^2})$.

Regarding computer sampling, you will have to look at computers with specifications that differ from the computer you are looking to purchase, called C_0. In other words, $D_c = \{(y_i, X_i) | X_i \neq X_0\}$ where x_0 are the specifications of your desired computer $X_i = (x_{i,1}, x_{i,2}, x_{i,3})$. In terms of computation, you will linearly regress $y_{i,1}$ on $x_{i,1}, x_{i,2}, x_{i,3}$. Employing such a method will allow you to construct a distribution of prices to offer using $o \sim \mathcal{N}\left(\hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3,\ \frac{1}{n-k-1} \sum_{i=1}^{n} \hat{\varepsilon}_i^2\right)$.

Regarding social sampling, you can observe the prices that other people have offered the clerk for the laptop you are interested in purchasing. In other words, $D_s = \{(y_i, X_i)X_i = X_0\}$. Unfortunately, you are not able to see whether the clerk accepted or rejected the offer. In terms of computation, you will compute the mean of the prices offered by others: $\frac{1}{n}\sum_{i=1}^{n} y_{i,s} = \bar{y}_s$. Side note: this estimator is consistent (sample is i.i.d. by definition so true by law of large numbers) and unbiased. Employing such a method will allow you to construct a distribution of prices to offer using $o \sim \mathcal{N}\left(\frac{1}{n}\sum_{i=1}^{n} y_i, \mathrm{SEM}^2\right)$.

Being human, you inherit the three following properties:

1. You would like to pay as little as possible for your laptop
2. You are averse towards performing cognitive labour that is highly computational
3. You would like your offer to be accepted by the clerk

These three characteristics are translated into the following utility function: $U(o, n_c, A(o)) = -\alpha o - \beta n_c + \gamma A(o)$ where:

1. α captures your preference to pay as little as possible for your laptop
2. β captures your aversion towards performing highly demanding cognitive labour (i.e. performing a multiple linear regression manually)
3. γ reflects how much you care about your offer being accepted by the clerk

Solving for this utility function, we are interested in:

$$\max_{o} \; U(o, n_c, A(o)),$$

where

$$A(o) = P(o > \tau | o) = \Phi\left(\frac{o - \mu_\tau}{\sigma_\tau}\right),$$

$$\frac{\partial U}{\partial o} = -\alpha + \gamma\,\phi\left(\frac{o - \mu_\tau}{\sigma_\tau}\right) \cdot \frac{1}{\sigma_\tau}$$

$$\frac{\partial U}{\partial o} = -\alpha + \gamma\,\frac{1}{\sigma_\tau}\phi\left(\frac{o - \mu_\tau}{\sigma_\tau}\right)$$

$$0 = -\alpha + \gamma\,\frac{1}{\sigma_\tau}\phi\left(\frac{o - \mu_\tau}{\sigma_\tau}\right)$$

$$\frac{\alpha\sigma_\tau}{\gamma} = \phi(z)$$

$$\frac{\alpha\sigma_\tau}{\gamma} = \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}$$

$$\pm\sqrt{-\ln(\frac{\alpha\sigma_\tau\sqrt{2\pi}}{\gamma})(2)} = z$$

$$\mu_\tau \pm (\sigma_\tau)\sqrt{-\ln(\frac{\alpha\sigma_\tau\sqrt{2\pi}}{\gamma})(2)} = o$$

We will choose the largest value to ensure that the offer gets accepted:

$$\mu_\tau + (\sigma_\tau)\sqrt{-\ln(\frac{\alpha\sigma_\tau\sqrt{2\pi}}{\gamma})(2)} = o$$

Table of Contents

Research Question

Does altering an agent's weights with respect to cognitive cost, offer acceptance, and offering price render the social sampling method more attractive than the computer sampling method in terms of utility?

Table of Contents

I constructed 30 computer prices for each possible combination of specifications. There were 2 types of CPUs (Tier 1, Tier 2) with an upgrade costing 300\$, two types of RAM (4 GB, 8GB) with a per unit cost of 30\$, and two types of SSD size (128 GB, 256 GB) with a per unit cost of 2\$. Using this deterministic price scheme, I added a random error term to generate the 30 different prices for the same computer specification where $\epsilon \sim \mathcal{N}(\text{deterministic price}, (\text{deterministic price})(0.12))$. For the clerk acceptance, we applied the same error term to the deterministic value of the desired computer given its specifications.

In total, my overall sample size was $2 \times 2 \times 2 \times 30 = 240$. Therefore, my computer sample size was 30 whilst my social sample size was 210.

The laptop that my agent was looking to purchase had a type 2 CPU with 4 GB of RAM and 128 GB of SSD.

Regarding my α, β, γ parameters, I increased each of them ceteris paribus to determine if there were any systematic changes in the optimal price offered and total utility.

Social sampling:

$$\mu_s + (\sigma_s)\sqrt{-\ln(\frac{\alpha\sigma_s\sqrt{2\pi}}{\gamma})(2)} = o$$

Computer sampling:

$$\mu_c + (\sigma_c)\sqrt{-\ln(\frac{\alpha\sigma_c\sqrt{2\pi}}{\gamma})(2)} = o$$

The baseline variables for α, β, γ were 1, 2, and 50 respectively.

For α, I increased its value from 1 to 20 using steps of 1 ceteris paribus.

For β, I increased its value from 1 to 20 using steps of 1 ceteris paribus.

For γ, I increased its value from 50 to 200 using steps of 10 ceteris paribus.

Table of Contents

Results

As α increased, the social estimation method always yielded higher utility when there was an interior optimum (i.e. $\frac{\alpha\sigma_\tau\sqrt{2\pi}}{\gamma} \in (0,1]$) .

As β increased, the social estimation method always yielded higher utility when there was an interior optimum.

As γ increased, the social estimation method yielded higher utility when there was an interior optimum.

Table of Contents

Limitations

External validity?

How should we derive my γ, β, α parameters in a real-world context?

Does this exercise, by construction, already answer my research question in the direction I expect?

What if the social price samples systematically deviated from the computer pricing samples?

How should I interpret these interior optimum issues?