Discrimination, Design, and Behavior in Secondary Markets: Evidence from a Natural Experiment on Vinted.cz

Tereza Burýšková*

CERGE-EI**

Abstract

JEL classification: D83, L82
Keywords: discrimination, online markets, rational inattention, platform design

*Email: tereza.buryskova@cerge-ei.cz
**A joint workplace of Charles University and the Economics Institute of the Czech Academy of Sciences, Politickych veznu 7, 11121 Prague, Czech Republic.

1 Introduction

2 Vinted.cz

3 Theoretical model

3.1 Two groups of sellers

Assume that there are two groups selling items on the platform. To highlight the relation to the context of Vinted.cz, I denote these by CZ and PL, and the proportion of PL sellers on the platform by p. These groups differ in the distribution of the quality of items they offer. The mean is equal for the two groups; however, the variance is different. I assume that the variance is larger for PL users. For simplicity, I assume that quality is distributed as

$$q_{PL} \sim \mathcal{N}(0,\sigma_{PL}^2), \tag{1}$$

$$q_{CZ} \sim \mathcal{N}(0,1), \tag{2}$$

where $\sigma_{PL}^2 \geq 1$.

3.2 Buyers

The buyer faces an item and decides

1. Whether or not to acquire more information I about the item (with potential cost of information).
2. Whether or not to buy it - $Y \in \{1,0\}$

If the buyer buys the item, she receives a payoff

$$U(q|Y=1,I) - C(I) \tag{3}$$

I assume that the user is risk-averse and derives exponential utility from the quality of the item. The expected utility of buying an item from a CZ and PL seller is therefore

$$U_{CZ} = -e^{\frac{\gamma^2}{2}}, \tag{4}$$

$$U_{PL} = -e^{\gamma^2 \frac{\sigma_{PL}^2}{2}}, \tag{5}$$

where $\gamma \in \mathbf{R}_0^+$ measures the risk-aversion of the agent (the higher, the more risk-averse). If the buyer does not buy the item, she gets a utility equal to $-\epsilon$, where $-\epsilon \in (U_{PL}, U_{CZ})$. This means that under perfect information, the user always wants to buy from a CZ seller, but never wants to buy from a PL seller.

Before buying, the user may be able to acquire additional information about the type of seller I. The agent chooses which information to acquire and faces a cost of information $C(I)$. Then, she forms a posterior belief about the agent being PL (denoted by r). After that, she takes an action $Y \in 1{,}0$ representing the decision to buy/not to buy. Then, the nationality of the seller is revealed, and the payoff is realized. The ex-ante problem of the agent, therefore, is

$$max_{Y,I} \quad Y \cdot [r \cdot U_{PL} + (1-r) \cdot U_{CZ} - C(I)] + (1-Y) \cdot [-\epsilon - C(I)] \,, \tag{6}$$

where r is the posterior belief about the seller being PL after consuming information I.

3.3 Platform

Since this version of the model abstracts from prices, I assume that the platform cares about the number of purchases (computed as the probability of buying). This could be a realistic proxy of the business model—the platform takes a fixed fee for every purchase.

3.4 Information structures

In this paper, I focus on four different information structures, each one related to a design of the platform's information policy.

3.4.1 Full costless info

In this case, I assume that the information about nationality is shown to the buyer without any searching $C(I) = 0$. There, the problem becomes quite simple—the posterior belief becomes 0 or 1, and buyers always buy from a CZ seller and never from a PL seller.

3.4.2 No info

In the opposite case, I consider the situation when the agent cannot recover the information about the seller's nationality. She is left only with her prior belief p.

3.4.3 Perfect costly signal

Then, I combine the two previous approaches and consider a case where a signal that completely reveals the true state of the world is available. However, this signal comes with a cost $C(I) > 0$. I define this cost to be proportional to the reduction in entropy between prior and posterior. That means that for an agent with prior p, the cost of this signal would be $C(I) = \lambda(-p \cdot log(p) - (1-p) \cdot log(1-p))$. Here, the decision of the agent is two-staged; first, she considers whether or not to acquire this signal; second, given the signal, she chooses to buy or not to buy.

3.4.4 Rational inattention

Finally, I consider a rationally inattentive buyer, who can flexibly acquire any information. The cost $C(I)$ of the information is proportional to the expected reduction in entropy after consuming the signal. After the signal is acquired, the agent chooses whether to buy or not to buy to maximize her payoff.

4 Solving the model

In this section, I describe the impact of information structures on buyers' behavior and also the platform's profits. I focus on four main outcomes

A Expected buyer payoff $\mathbf{E}[U]$

B Probability of buying $P(Y=1)$

C Behavioral discrimination, which I define as $P(Y=1|CZ) - P(Y=1|PL)$, representing the difference between the probability of buying from a CZ seller and a PL seller.

D Amount of info acquired, which I represent by $H(p) - H(r)$ - the reduction in entropy between prior and posterior belief.

4.1 Full info

In this case, the computation is straightforward. The buyer always takes the *correct* action after learning the nationality of the seller. Therefore

$$\mathbf{E}_{FI}[U] = -p\epsilon + (1-p)E_{CZ} \tag{7}$$

$$P_{FI}(Y=1) = 1-p \tag{8}$$

$$P(Y=1|CZ) - P(Y=1|PL)_{FI} = 1 \tag{9}$$

$$H(p) - H(r)_{FI} = -p \cdot log(p) - (1-p)log(1-p), \tag{10}$$

where $H(x) = -x \cdot log(x) - (1-x)log(1-x$

4.2 No info

Here, the agent is left only with her prior. Therefore, we need to consider each combination of buying decision and the nationality of the user separately

$$\mathbf{E}_{NI}[U] = \max((1-p) \cdot U_{CZ} + p \cdot U_{PL}, -\epsilon) \tag{11}$$

$$P_{NI}(Y=1) = \mathbf{I}(p \cdot U_{PL} + (1-p) \cdot U_{CZ} > -\epsilon) \tag{12}$$

$$P(Y=1|CZ) - P(Y=1|PL)_{NI} = 0 \tag{13}$$

$$H(p) - H(r)_{NI} = 0 \tag{14}$$

4.3 Perfect costly signal

Here, the agent first chooses whether or not acquire the signal (which is fully given by the parameters of the model), and then chooses the optimal action Y. For clarity, I denote by S a dummy indicating whether the signal was acquired or not. This can be expressed as

$$S = \mathbf{I}\left\{-\lambda H(p) + \mathbf{E}_{FI}[U] > \mathbf{E}_{NO}[U]\right\} \tag{15}$$

$$\mathbf{E}_{PS}[U] = \max(-\lambda H(p) + \mathbf{E}_{FI}[U], \mathbf{E}_{NI}[U]) \tag{16}$$

$$P_{PS}(Y=1) = S \cdot (1-p) + (1-S) \cdot \mathbf{I}\left[p \cdot U_{PL} + (1-p)E_{CZ} > -\epsilon\right] \tag{17}$$

$$P(Y=1|CZ) - P(Y=1|PL)_{PS} = S \tag{18}$$

$$H(p) - H(r)_{PS} = S \cdot H(p) \tag{19}$$

4.4 Rational inattention

Finally, I discuss the most complicated case of rational inattention. There, the agent chooses information flexibly to maximize her objective function, which represents the expected payoff minus the cost of information measured by mutual information expressed in terms of the probabilities

$$\begin{aligned}\mathbf{E}_{RI}\left[U\right] = max_{a,b} p \cdot \left[aU_{PL} - (1-a)\cdot\epsilon\right] + (1-p)\cdot\left[b\cdot E_{CZ} - (1-b)\cdot\epsilon\right] - \\ - \sum_{x\in\{CZ,PL\},Y} P(x,y)\cdot log(P(x,y)/P(x)P(y)\end{aligned} \tag{20}$$

where the last term could be simplified in our simple binary setting to

$$\begin{aligned}\mathbf{E}_{RI}\left[U\right] = \max_{a,b\in[0,1]} \quad & p\cdot\left[aU_{PL} - (1-a)\cdot\epsilon\right] + (1-p)\cdot\left[b\cdot E_{CZ} - (1-b)\cdot\epsilon\right] \\ & - pa\cdot\log\left(\frac{a}{pa+(1-p)b}\right) + p(1-a)\cdot\log\left(\frac{1-a}{p(1-a)+(1-p)(1-b)}\right) \\ & + (1-p)b\cdot\log\left(\frac{b}{pa+(1-p)b}\right) + (1-p)(1-b)\cdot\log\left(\frac{1-b}{p(1-a)+(1-p)(1-b)}\right).\end{aligned} \tag{21}$$

Unfortunately, this problem does not have a closed-form solution. Therefore, to determine the expected utility, I compute the solution numerically. I denote the optimal values of a and b by a^* and b^*. These values have a straightforward interpretation - they represent the conditional probabilities of buying in an optimum

$$a^* = P\left[Y=1|PL\right], \tag{22}$$

$$b^* = P\left[Y=1|CZ\right]. \tag{23}$$

Once we have these values, the computation of other parameters of interest is straightforward

$$P_{RI}(Y=1) = a^*\cdot p + (1-p)\cdot b^* \tag{24}$$

$$P(Y=1|CZ) - P(Y=1|PL)_{PS} = b^* - a^* \tag{25}$$

$$
\begin{aligned}
H(p) - H(r)_{RI} = &- pa^* \cdot \log\left(\frac{a^*}{pa^* + (1-p)b^*}\right) \\
&+ p(1-a^*) \cdot \log\left(\frac{1-a^*}{p(1-a^*) + (1-p)(1-b^*)}\right) \\
&+ (1-p)b^* \cdot \log\left(\frac{b^*}{pa^* + (1-p)b^*}\right) \\
&+ (1-p)(1-b^*) \cdot \log\left(\frac{1-b^*}{p(1-a^*) + (1-p)(1-b^*)}\right).
\end{aligned} \tag{26}
$$

5 Predictions