Optimal Menu when Agents Make Mistakes*

Sergei Mikhalishchev†(CERGE-EI)‡

Abstract

This paper studies how optimal menu chosen by a social planner depend on whether agents receive imperfect signal about her true taste (imperfect self-knowledge) or the properties of available alternatives (imperfect information). Under imperfect self-knowledge it is not optimal to offer fewer alternatives than the number of different tastes present in the population unless noise is infinite (agents have no clue about their true preferences). As noise increases, the social planner would offer menu items closer together (more similar), in the limit only offering one choice matching the mean preference in the population. However, under imperfect information, as noise increases the social planner wants to restrict the number of alternatives. Whether he makes them more or less similar is non-linear in noise.

Keywords: Discrete choice, Optimal menu, Bounded rationality, Welfare analyses

JEL classification codes: D30, D60, D81, H80

---

*This project received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 678081) and institutional support RVO 67985998 from the Czech Academy of Sciences. I thank Ole Jann, Marco Mariotti, Andrei Matveenko, Vladimır Novak, Avner Shaked, Jakub Steiner, and Ekaterina Travova for their comments and suggestions. I also thank Jan Zapal for continual support on this project and Filip Matějka for excellent supervision and patient guidance. All remaining errors are my own.

†e-mail: sergei.mikhalishchev@cerge-ei.cz

‡CERGE-EI, a joint workplace of Charles University and the Economics Institute of the Czech Academy of Sciences, Politickych veznu 7, 111 21 Prague, Czech Republic.

1 Introduction

In life, we often face choice from a discrete menu, for example when choosing an insurance plan, school for our children, or pension fund. At the same time, when confronted with these important decisions we often make mistakes, for two potential reasons. First, we misperceive the true properties of alternatives, i.e. we have imperfect information. For example, individuals are uninformed and underestimate potential cost savings from changing prescription drug plans (Kling et al., 2012), not fully informed about crucial aspects of an insurance plans (Handel and Kolstad, 2015), and when choosing a car think of fuel costs as scaling linearly in miles per gallon instead of gallons per mile (Allcott, 2013). Second, we misperceive our own tastes, i.e. we have imperfect self-knowledge. For example, individuals overestimate their attendance as well as their likelihood of cancelling automatically renewed memberships when choosing a gym contract (DellaVigna and Malmendier, 2006), and more generally are myopic in their decisions, may lack skill in predicting their tastes and risk preference, and can be led to erroneous choices by fallible memory and incorrect evaluation of past experiences (Kahneman, 1994; Heckman et al., 2021).

In the examples above, a government or other social planner could regulate the size of the menu from which consumers choose, as well as the properties of alternatives within it. The social planner can not possibly know the individual taste of a particular agent and, hence, is not able to provide the first best alternative for each agent. However, knowing the overall population characteristics, including probabilities of mistakes and distribution of tastes, he can construct a menu of alternatives, referred to as *optimal menu*, that maximizes the sum of expected utilities of agents.

I analyze an optimal menu under assumptions that agents misperceive either the true properties of available alternatives or their own tastes. In two limiting cases, when the misperception is insignificant or consumers pick the alternative randomly, the optimum menu is identical under both type of misperception. For intermediate degrees of rationality, the dependence of the optimum choice set on the precision of choice is complex. We use simple setting and numerical calculations and demonstrate that when agents misperceive the available options it is optimal to limit choice when the probabilities of mistakes are moderately high. Additionally, it could be optimal to construct a menu with more distinct alternatives. In contrast, when agents misperceive their own tastes, it is optimal to limit choice solely when agents choose randomly, and to propose alternatives that are more similar when there is a greater probability of a mistake.

The intuition behind the results is that when agents misperceive the properties of alternatives, every additional alternative in the menu has the benefit of providing more choice (matching the agents' taste more precisely) at the cost of increasing probability and magnitude of mistakes. Thus, the more similar the alternatives, the more difficult it is for the agent to differentiate between them. Therefore, it could be optimal to construct a menu with more distinct alternatives in order to decrease the probability of a mistake, depending on the distribution of tastes in the population. At the same time, when the probability of a mistake is large it becomes optimal to remove the options that could induce large utility loss and leave one option that matches the mean taste in the population.

In contrast, when agents have imperfect self-knowledge, the probability of a mistake depends only on the midpoints between properties of alternatives. Thus, the probability of a mistake would not be decreased if alternatives were differentiated. Moreover, since the probability of mistakes linearly affected by alternatives, it is weakly beneficial to introduce more alternatives in the menu.

The discussion about individuals misperceiving the true properties of alternatives and accordingly failing to choose the best one goes back at least as far as Luce (1959), who analyzes agent choice subject to a random noise. Mirrlees (1987, 2017) and Sheshinski (2003b,a, 2010, 2016) study the welfare maximization problem when agents misperceive the true properties of alternatives. They show that while, the choice should not be limited when the agents are completely rational, the optimum choice-set is a singleton when the probability of a mistake is relatively high. In contrast, this paper focuses on comparing optimal menu allocations in two situations:

when the agent misperceives either the true properties of alternatives or her own taste.

In recent years, a growing literature in industrial organization has analyzed the situation when a firm interacts with boundedly rational agents. For a classic textbook treatment see Anderson et al. (1992); more recent papers include Kamenica (2008), Hefti (2018), Persson (2018), and Gerasimou and Papi (2018). A review of other studies on complexity and manipulation can be found in Spiegler (2016). The main focus of this literature is on the market environment, and the agents' limitations arise solely from misperception of the true properties of available alternatives. This study considers two sources of mistakes and focuses on the welfare maximization problem.

In addition, this paper proposes a new explanatory insight into the choice paradox (Schwartz, 2004), i.e. the effect when a larger choice set sometimes decreases the satisfaction of individuals and even leads to the rejection of the offer. This phenomenon has been observed, for example, when consumers purchased jam and chocolate (Iyengar and Lepper, 2001) as well as when they made more important decisions such as the choice of 401k pension plans (Iyengar et al., 2004) or participation in an election (Nagler, 2015)[1]. At the same time, several studies suggest that the existence of the choice paradox and the efficiency of corresponding interventions, such as categorization of goods, depend on whether consumers are familiar with the product or not (Chernev, 2003; Mogilner et al., 2008). There are numerous models that attempt to explain this evidence (Irons and Hepburn, 2007; Sarver, 2008; Ortoleva, 2013; Kuksov and Villas-Boas, 2010). While my study does not focus on a particular mechanism, it suggests that the existence of this phenomenon and relevant interventions depend on the source of mistakes in the decision making process. Thus, when agents misperceive the true properties of alternatives we can observe choice overload, and limiting the menu size could be a welfare maximizing intervention. However, when agents have imperfect self-

[1]Further discussion on empirical evidence when choice is bad can be found, for example, in Scheibehenne et al. (2010) or Chernev et al. (2015).

knowledge, we would not observe the choice overload and, hence, should not limit the choice.

The rest of the paper is organized as follows. The next section presents the model setup. Section 3 discusses a simple model with two agents in order to illustrate the intuition behind the results, and then provides numerical simulations with populations of agents. The last section concludes.

2 Model

A population of $M \geq 2$ agents chooses from a set of $N \geq 2$ alternatives. The utility of the agent $i \in \{1, ..., M\}$ from the alternative $j \in \{1, ..., N\}$ is $U_i^j = -(t_i - v^j)^2$, where $t_i \in \mathbb{R}$ is the taste (bliss point) of i and $v^j \in \mathbb{R}$ is the property of j. $T \geq 2$ is the number of unique tastes in the population. The agent misperceives parameters of the model. I describe two versions of the model:

– **with misperceived true properties of alternatives:** the agent observes a signal $\vartheta_i^j = v^j + e_i^j$, where v^j is a true property of the option, and noise e_i^j is a random variable drawn from the distribution with mean zero and variance σ_i^j. She chooses the alternative with the signal that is a closest match to her taste[2], i.e. solves the following problem:

$$\max_{j \in \{1,...,N\}} -(t_i - \vartheta^j)^2.$$

– **with misperceived own true taste:** the agent observes a signal $\tau_i = t_i + e_i$, where t_i is the true taste of the agent, and noise e_i is a random variable drawn from the distribution with mean zero and variance σ_i. She chooses the alternative with the property that is a closest match to the signal of her taste, i.e. solves the following problem:

$$\max_{j \in \{1,...,N\}} -(\tau_i - v^j)^2.$$

In both versions of the model if there are several alternatives that solve the

[2]For discussion on when this behavior is optimal for the agent see Weibull et al. (2007).

agent's problem then the agent chooses randomly between them.

The social planner maximizes overall welfare by choosing a number and properties of available alternatives, i.e. the optimal menu:

$$\max_{N, v^j \forall j \in \{1,\ldots,N\}} \sum_{i=1}^{M} \sum_{j=1}^{N} P_i^j U_i^j ,$$

where P_i^j is the probability that the agent i chooses the option j. I assume that $N \leq T$: the maximum number of options that the social planner could propose is equal to the number of tastes in the population.[3]

The problem has the following time-line:

1. The social planner observes (i) distributions of mistakes, and (ii) what the tastes in the population are, and (iii) the number of agents with each taste.
2. He chooses the optimal menu.
3. Agents observe signals.
4. They choose an alternative from the menu.

3 Solution

The solution to the welfare maximization problem depends on the size of the noise. Regardless of the source of mistakes, when there is no noise the social planner creates a menu with alternatives that match tastes perfectly; when noise is infinite, it is optimal to limit choice and provide only one alternative that matches the mean taste in the population. This result is formalized in Propositions 1 and 2.

Proposition 1. *If $\sigma_i^j = 0$ or $\sigma_i = 0 \; \forall (i,j)$, then $N = T$, $v^j = t_i$.*

Proof. Since $U_i \leq 0 \; \forall i \Rightarrow \max(\sum_{i=1}^{M} \sum_{j=1}^{N} P_i^j U_i^j) = 0$ which is obtained when $N = T$, $v^j = t_i$. □

[3] I make this assumption because the welfare function is not monotone in the number of options: for example, if for a given distribution the optimal number of alternatives is 4 then the solution to the welfare maximization problem automatically includes any number that is divisible by 4.

Proposition 2. *If* $\sigma_i^j \to \infty$ *or* $\sigma_i \to \infty$ $\forall (i,j)$*, then* $N = 1$ *and* $v^j = \frac{\sum t_i}{M}$.

Proof. If $\sigma_i^j \to \infty$ or $\sigma_i \to \infty$, then all alternatives are a priori the same for agents $P_i^j = \frac{1}{N}$. The solution to the welfare maximization problem is $N = 1$ and $v^j = \frac{\sum t_i}{M}$. □

In the next subsection I illustrate the solution of the model for the intermediate cases using a model with uniformly distributed noise and two agents. Then I show that the results obtained are valid for the larger population of agents with continuous distribution of noise using numerical simulations.

3.1 Two agents

There are two agents, $i \in \{1, 2\}$, with tastes symmetrically allocated around zero, $t_1 = -t_2 < 0$.[4] The social planner could propose at most two options, $j \in \{1, 2\}$. I assume that $v^1 \leq v^2$. The situation when $v^1 = v^2$ is identical to the situation when the social planner proposes only one alternative and limits the agents' choice.

I assume that the noise is uniformly distributed, e_i^j and $e_i \sim U(-b, +b)$. Therefore, the social planner expects that agent 1 chooses the first option with probability P_1^1 and the second option with probability P_1^2. Agent 2 chooses similarly.

In the case of misperceived true properties of alternatives, the probabilities are as follows:

$$P_1^1 = \min\left(1, \max\left(0,\ 1 - 0.5 * (\tfrac{v^1 - v^2 + 2b)}{2b})^2\right)\right), \quad P_1^2 = 1 - P_1^1.$$
$$P_2^1 = \min\left(1, \max\left(0,\ 0.5 * (\tfrac{v^1 - v^2 + 2b)}{2b})^2\right)\right), \quad P_2^2 = 1 - P_2^1.$$

In the case of misperceived true own tastes, the probabilities are as follows:

$$P_1^1 = \min\left(1, \max\left(0, \frac{\frac{v^1+v^2}{2} - (t_1 - b)}{2b}\right)\right), \quad P_1^2 = 1 - P_1^1.$$
$$P_2^1 = \min\left(1, \max\left(0, \frac{\frac{v^1+v^2}{2} - (t_2 - b)}{2b}\right)\right), \quad P_2^2 = 1 - P_2^1.$$

[4] It is without loss of generality since for any two distinct tastes one always can re-scale it to be symmetrically allocated around zero.

In order to obtain an analytical solution for both cases I have to make an additional heroic assumption that $v^1 = -v^2$.[5] This symmetry assumption simplifies the characterization of the solution.[6] The solution to the welfare maximization problem is formalized in Propositions 3 and 4.

Proposition 3. *In the case of misperceived true values of alternatives, the welfare maximization problem has the following solution:*

– ***small noise*** *($b \leq |t_i|$): $v^1 = -v^2 = t_1$;*
– ***medium noise*** *($|t_i| < b < 4|t_i|$): $v^1 = -v^2 = \frac{-b^2-4bt_1}{3t_1}$;*
– ***large noise*** *($4|t_i| \leq b$): $v^1 = v^2 = 0$.*

Proof. See Appendix A1. □

Proposition 4. *In the case of misperceived true own tastes, the welfare maximization problem has the following solution:*

– ***small noise*** *($b \leq |t_i|$): $v^1 = -v^2 = t_1$;*
– ***medium and large noise*** *($|t_i| < b$): $v^1 = -v^2 = -\frac{t_1^2}{b}$.*

Proof. See Appendix A2. □

Accordingly, when the noise is small ($b \leq |t_i|$), in both cases the social planner proposes options that match the tastes of the agents perfectly, and they choose the option closest to their true taste with certainty. When the noise is significantly large ($|t_i| < b$), then the solution depends on the source of mistakes. If agents misperceive the true properties of alternatives, it is optimal to limit the choice when the noise is finitely large. However, when agents misperceive their tastes, it is optimal to propose two alternatives with different properties for any finite noise.

In addition, if agents misperceive the true properties of alternatives, there exists noise ($|t_i| < b < 2|t_i|$) when the difference in the properties of proposed alternatives increases in the noise, i.e. $\frac{\partial v^1}{\partial b} < 0$ and $\frac{\partial v^2}{\partial b} > 0$. However, if agents misperceive

[5] It also could be interpreted as if the welfare function satisfies Ralwsian principle of social justice, i.e. overall welfare is based on the welfare of society's worse-off member.

[6] Without a symmetry assumption the solution for the situation when the agent misperceives the true properties of alternatives would be asymmetrical and (possibly) not unique. While the solution for the situation when the agent has imperfect self-knowledge is the same.

their tastes, the social planner always proposes alternatives that are more similar the larger the noise is.

3.1.1 Intuition

The results are driven by the fact that if a taste is unclear, the distance between true taste and the properties of the options is distorted in the same way for all options, while if the properties of the options are unclear this distortion is different for any option.

In particular, let's denote t_1 as t and v^1 as v and rewrite the probability that the agent makes the wrong choice (i.e. she chooses the alternative that is not the closest to her true taste) as follows. In the case of misperceived true properties of alternatives:

$$P_1^2 = P_2^1 = 0.5(\frac{2v + 2b}{2b})^2.$$

In the case of misperceived true own tastes:

$$P_1^2 = P_2^1 = \frac{\frac{v+(-v)}{2} - t + b}{2b}.$$

Therefore, when the noise originates from the misperception of alternatives, placing options close to each other increases the probability of a mistake, which is a nonlinear function of v. Thus, there is an inverted U-shaped curvilinear relationship between the optimal property of the alternative v and size of the noise, as depicted in Figure 1. Thus, when the noise is significant, but still small ($|t_i| < b < 2|t_i|$), the social planner wants to distance the properties of alternatives from each other. In this situation, the loss from the decrease in utility in the case of the correct choice is smaller than the gain from the decrease in the probability of the wrong choice. However, when the noise is moderately large ($2|t_i| \leq b < 4|t_i|$), it is not profitable to distance the properties of alternatives further away from each other. The loss from the decrease in utility in the case of the correct choice outweighs the gain from the decrease in the probability of the wrong choice. Therefore, the social planer chooses properties of alternatives closer to each other. When the probability of the wrong choice is significantly high ($4|t_i| \leq b$), it is optimal to propose alternatives with identical properties.

However, when agents misperceive their tastes, the probability of a mistake depends linearly on the midpoint between properties of alternatives. Therefore, it is not beneficial to differentiate properties of alternatives, since doing so does not decrease the probability of the wrong choice. Accordingly, the social planner chooses v by equalizing the marginal gain of locating an option closer to the center for the second agent (reducing the loss in the case of the wrong choice) and marginal loss for the first agent (reducing the gain in the case of the correct choice) given the probabilities of mistakes.

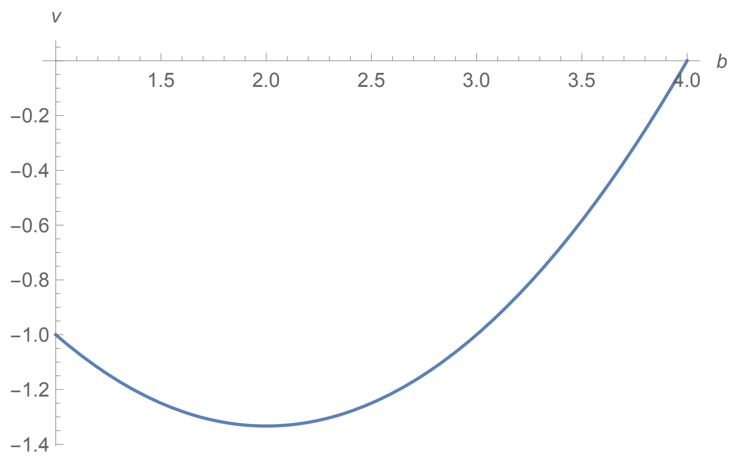

Figure 1: Optimal property of the first alternative as a function of b and $t_1 = -1$.

3.2 Many agents

3.2.1 Setup

There is a single-peaked population of agents with the variety of tastes $T = 7$. When agents misperceive the true properties of alternatives, e_i^j is assumed to be identically and independently Gumbel distributed. The Gumbel distribution has fatter tails than a Normal distribution; however, the difference between them is often indistinguishable empirically (Train, 2002). At the same time, the difference of Gumbel distributed variables, which is used for calculating probabilities of agents' choice, follows the Logistic distribution. This significantly simplifies the numerical simulation. Therefore, the probability that agent i chooses option j is:

$$P_i^j = \frac{\exp(U_i^j/\lambda)}{\sum_i^N \exp(U_i^j/\lambda)}.$$

When agents misperceive their own true tastes, e_i is assumed to be identically and independently Logistic distributed.[7] In this case, the probability that agent i chooses option j is:

$$P_i^j = \int_{\frac{v^{j-1}+v^j}{2}}^{\frac{v^j+v^{j+1}}{2}} \frac{\exp(\frac{t_i-v^j}{0.5\lambda})}{0.5\lambda(1+\exp(\frac{t_i-v^j}{0.5\lambda}))^2} dv^j.$$

In both situations higher values of λ correspond to larger variance and, hence, a higher probability of making a mistake. I solve for every possible menu size and then select the one that maximizes welfare.[8]

3.2.2 Results

The solution with the optimal number of alternatives and optimal menu allocation is presented in Figures 2-5 for different λ. The grey bars (histogram) correspond to the number of agents with a particular taste. The optimal properties of alternatives are defined by vertical lines. The optimal number of options is stated above the graphs. In some situations there are fewer vertical lines than the optimal number of alternatives, since there are several identical options that match the same taste. Intuitively, additional options with repeated values increase the probability that agents choose a particular alternative. Thus, when one taste is more salient in the

[7] In this case, I do not use the Gumbel distribution, since it is asymmetric. In turn, the asymmetry property skews the optimal menu, complicating the visual comparison. However, the qualitative results of the welfare analysis with the Gumbel distribution are identical to the analysis with the Logistic distribution.

[8] Calculations are performed in R using the "optimx" package.

population, it is beneficial to highlight the alternative that matches this taste.[9,10]

Figure 2 shows that when the noise is small, it is optimal to provide alternatives that match tastes perfectly under both kinds of mistakes.

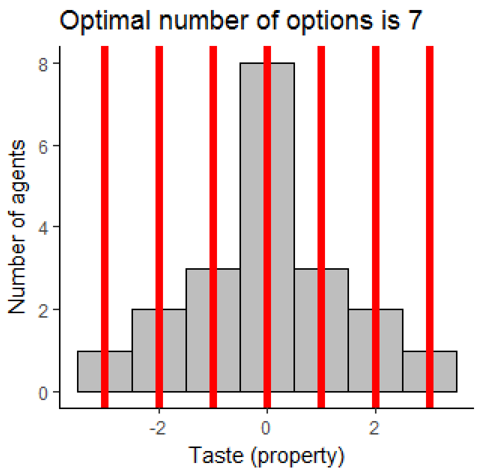

Figure 2: Optimal menu allocation when agents misperceive the true properties of alternatives or their own tastes, and $\lambda = 0.1$. The red lines indicate the optimal properties of alternatives. The histogram shows the distribution of agents.

Figures 3 and 4 show the optimal menus for the situation when the noise is significantly large. When agents misperceive true properties of alternatives, it is optimal to limit choice (Figures 3). When the probabilities of mistakes increase, the social planner decreases the menu size. At the same time, when agents misperceive their own taste, it is not optimal to limit choice (Figures 4). Thus, the social planner proposes 7 alternatives with unique properties for any noise. When the probabilities of mistakes increase, he allocates alternatives closer to each other and to the mean taste in the population.

It is worth noticing that the effect of the decrease in the tastes' inequality is

[9] Mirrlees (2017) refers to such manipulation as "advertising". One possible type of "advertising" is nudges. For example, it was shown that setting an option as a default increases the probability that this alternative will be chosen. See Thaler and Sunstein (2008) for additional discussion on the topic.

[10] One way to avoid the presence of identical options in the menu is to introduce the following probability function: $P_i^j = \frac{m(j)P_i^j}{\int m(y)P_i^y dy}$, where $m(j)$ is a density of alternatives with identical properties (Mirrlees, 2017). This formula relates to the modified multinomial logit model by Matějka and McKay (2015). Accordingly, another possible explanation for the "advertising" effect is prior knowledge of agents about options in a menu.

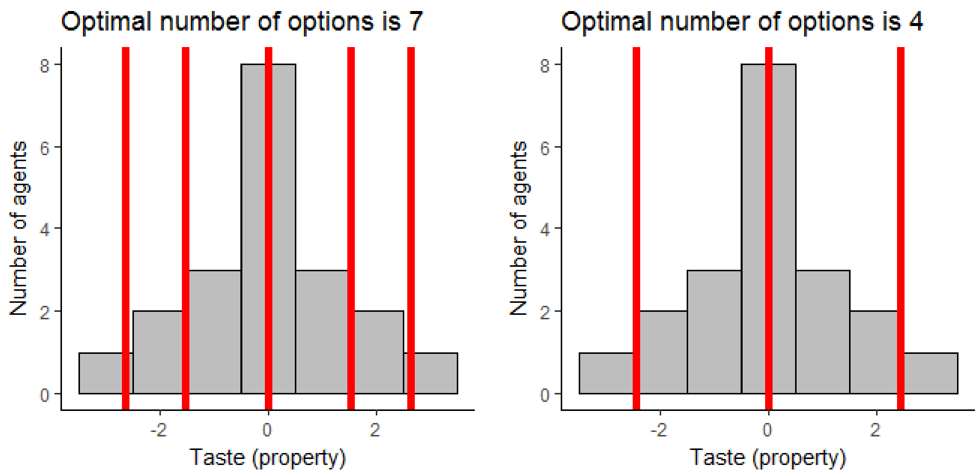

Figure 3: Optimal menu allocation when agents misperceive the true properties of alternatives for different noise ($\lambda = 1$ on the left and $\lambda = 2$ on the right graph). The red lines indicate the optimal properties of alternatives. The histogram shows the distribution of agents.

similar to the decrease in noise. Figure 5 shows the optimal menu allocation for the different population of agents with the same variety of tastes $T = 7$, but with lower density of agents with the most frequent (mode) taste $t_{mode} = 0$. In this situation, when agents misperceive the true properties of alternatives (left graph, Figure 5), the social planner proposes more alternatives to agents, compared to the optimal menu for a population with higher density of agents with mode taste (left graph, Figure 3). Similarly, when agents misperceive their own taste (right graph, Figure 5), the social planner proposes 7 alternatives, but allocates them further away from each other and the mean taste in the population, compared to the optimal menu for a population with a higher density of agents with mode taste (left graph, Figure 4).

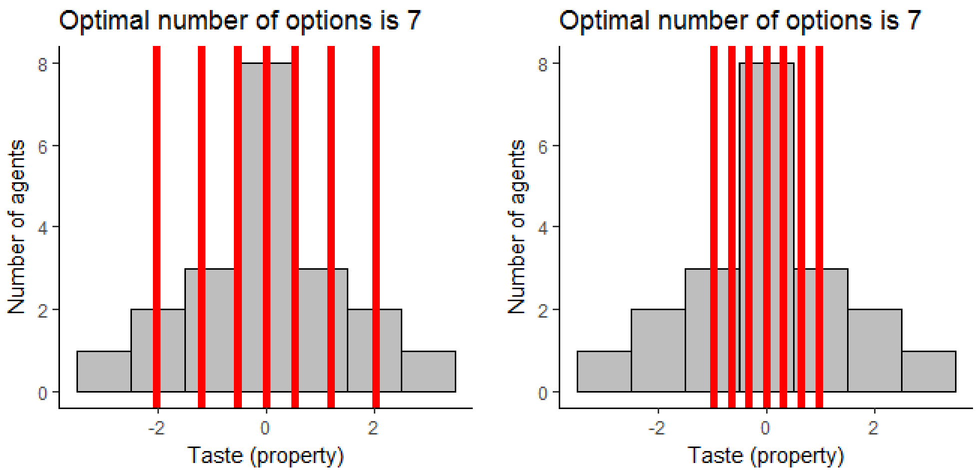

Figure 4: Optimal menu allocation when agents misperceive their own tastes for different noise ($\lambda = 1$ on the left and $\lambda = 2$ on the right graph). The red lines indicate the optimal properties of alternatives. The histogram shows the distribution of agents.

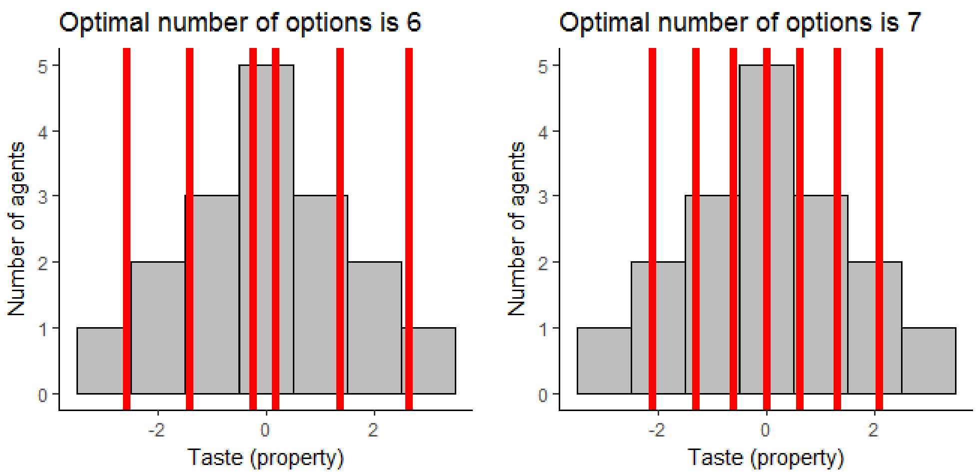

Figure 5: Optimal menu allocation when agents misperceive the true properties of alternatives (left graph) or their own tastes (right graph) and $\lambda = 1$. The red lines indicate the optimal properties of alternatives. The histogram shows the distribution of agents.

4 Conclusion

Although there is a large body of literature that studies problems with agents who make mistakes, there is still a lack of studies that analyze a discrete choice problem with heterogeneous agents and the social planner. This paper provides solution to the welfare maximization problem and shows that if agents misperceive the true properties of alternatives, the optimal menu differs significantly from one when agents misperceive their own tastes. Therefore, this study suggests that when designing a menu set one should take into account not only the demand for a particular alternative, but also the probability and source of a mistake.

References

Allcott, H. (2013). The welfare effects of misperceived product costs: Data and calibrations from the automobile market. *American Economic Journal: Economic Policy 5*(3), 30–66.

Anderson, P. S., A. de Palma, and J.-F. Thisse (1992). *Discrete Choice Theory of Product Differentiation.* Cambridge, MA: MIT Press.

Chernev, A. (2003). Product assortment and individual decision processes. *Journal of Personality and Social Psychology 85*(1), 151–162.

Chernev, A., U. Bockenholt, and J. Goodman (2015). Choice overload: Conceptual review and meta-analysis. *Journal of Consumer Psychology 25*(2), 333–358.

DellaVigna, S. and U. Malmendier (2006). Paying not to go to the gym. *American Economic Review 96*(3), 694–719.

Gerasimou, G. and M. Papi (2018). Duopolistic competition with choice-overloaded consumers. *European Economic Review 101*, 330–353.

Handel, B. R. and T. J. Kolstad (2015). Health insurance for "humans": Information frictions, plan choice, and consumer welfare. *American Economic Review 150*(8), 2449–2500.

Heckman, J. J., T. Jagelka, and T. Kautz (2021). Some contributions of economics to the study of personality.

Hefti, A. (2018). Limited attention, competition and welfare. *Journal of Economic Theory 178*, 318–359.

Irons, B. and C. Hepburn (2007). Regret theory and the tyranny of choice. *The Economic Records 86*(261), 191–203.

Iyengar, S. S., G. Huberman, and W. Jiang (2004). How much choice is too much? contributions to 401(k) retirement plans. In O. S. Mitchell and S. Utkus (Eds.), *Pension Design and Structure: New Lessons from Behavioral Finance*, pp. 83–95. Oxford University Press.

Iyengar, S. S. and M. Lepper (2001). When choice is demotivating: Can one desire too much of a good thing? *Journal of personality and social psychology 79*, 995–1006.

Kahneman, D. (1994). New challenges to the rationality assumption. *Journal of Institutional and Theoretical Economics 150*, 18–36.

Kamenica, E. (2008). Contextual inference in markets: On the informational content of product lines. *American Economic Review 98*(5), 2127–2149.

Kling, J. R., S. Mullainathan, E. Shafir, L. C. Vermeulen, and M. V. Wrobel (2012). Comparison friction: Experimental evidence from medicare drug plans. *The Quarterly Journal of Economics 127*(1), 199–235.

Kuksov, D. and J. M. Villas-Boas (2010). When more alternatives lead to less choice. *Marketing Science 29*(3), 507–524.

Luce, R. D. (1959). *Individual Choice Behavior: A Theoretical Analysis*. Wiley.

Matějka, F. and A. McKay (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review 150*(1), 272–298.

Mirrlees, A. J. (1987). Economic policy and nonrational behavior. UC Berkeley: Department of Economics, UCB.

Mirrlees, A. J. (2017). *Bounded Rationality and Economic Policy*. `https://rb.gy/opw1cj`. Accessed August 25, 2020.

Mogilner, C., T. Rudnick, and S. S. Iyengar (2008). The mere categorization effect: How the presence of categories increases choosers' perceptions of assortment variety and outcome satisfaction. *Journal of Consumer Research 35*(2), 202–215.

Nagler, M. G. (2015). Trading off the benefits and costs of choice: Evidence from australian elections. *Journal of Economic Behavior and Organization 114*, 1–12.

Ortoleva, P. (2013). The price of flexibility: Towards a theory of thinking aversion. *Journal of Economic Theory 148*(3), 903–934.

Persson, P. (2018). Attention manipulation and information overload. *Behavioural Public Policy 2*(1), 78–106.

Sarver, T. (2008). Anticipating regret: Why fewer options may be better. *Econometrica 76*(6), 263–305.

Scheibehenne, B., R. Greifeneder, and M. P. Todd (2010). Can there ever be too many options? a meta-analytic review of choice overload. *Journal of Consumer Research 37*(3), 409–425.

Schwartz, B. (2004). *The paradox of choice: Why more is less.* New York: Harper Collins.

Sheshinski, E. (2003a). Bounded rationality and socially optimal limits on choice in a self-selection model.

Sheshinski, E. (2003b). Optimal policy to influence individual choice probabilities.

Sheshinski, E. (2010). Limits on individual choice.

Sheshinski, E. (2016). A note on income taxation and occupational choice. *Research in Economics 70*(1), 20–23.

Spiegler, R. (2016). Choice complexity and market competition. *Annual Review of Economics 8*, 1–25.

Thaler, H. R. and R. C. Sunstein (2008). *Nudge: Improving decisions about health, wealth, and happiness.* Yale University Press.

Train, K. (2002). *Discrete Choice Methods With Simulation.* Cambridge University Press.

Weibull, J. W., L.-G. Mattsson, and M. Voorneveld (2007). Better may be worse: Some monotonicity results and paradoxes in discrete choice under uncertainty. *Theory and Decision 63*(2), 121–151.

A

Proof of Proposition 3

Because of the symmetry assumption, the welfare maximization problem could be reduced to the choice of one variable $v^1 = v \leq 0$. I denote $t_1 = t < 0$. If $b < |t|$, then the probability of a mistake equals zero and the first best allocation is optimal. Therefore, I consider a situation when $b \geq |t|$ and $0 \leq P_i^j \leq 1 \forall i, j$. Then the welfare maximization problem is the following:

$$\max_v W(v) = \Big\{(1 - 0.5(\frac{2v + 2b}{2b})^2) \cdot -2(t - v)^2 + 0.5(\frac{2v + 2b}{2b})^2 \cdot -2(t + v)^2\Big\}.$$

The derivative with respect to v is the following:

$$0.5(t - v)^2(2b + 2v) - 0.5(t + v)^2(2b + 2v) - \\ 0.25(t + v)(2b + 2v)^2 + 2(t - v)(b^2 - 0.125(2b + 2v)^2) = 0.$$

This equation has two solutions:

$$v = 0,$$

$$v = \frac{-b^2 - 4bt}{3t}.$$

Since $v \leq 0$, the second solution exists only for $b \leq 4|t|$. Moreover, when $b = 4|t|$, then $v = 0$ and the two solutions coincide. In this situation the welfare is $W(b = 4|t|) = -2t^2$. At the same time, if one substitutes $v = \frac{-b^2-4bt}{3t}$ into the maximization problem, then $W(b = |t|) = 0$ and $W > -2t^2$ for any $|t| < b < 4|t|$. Therefore, for $b < 4|t|$ the welfare is maximized when $v = \frac{-b^2-4bt}{3t}$; for $b \geq 4t$ it is optimal to provide the menu with two identical alternatives $v^1 = v^2 = 0$. □

B

Proof of Proposition 4

If the $b < |t|$, then the probability of a mistake equals zero and the first best allocation is optimal. Therefore, I consider a situation when $b \geq |t|$ and $0 \leq P_i^j \leq 1 \; \forall i, j$. Then the welfare maximization problem is the following:

$$\max_v \Big\{ \frac{-t+b}{2b} \cdot -(t-v)^2 + (1 - \frac{-t+b}{2b}) \cdot -(t+v)^2 + \frac{t+b}{2b} \cdot -(t+v)^2 + (1 - \frac{t+b}{2b}) \cdot -(t-v)^2 \Big\}.$$

The derivative with respect to v is the following:

$$-4(t^2 + bv) = 0.$$

Therefore, the solution to the welfare maximization problem is:

$$v = -\frac{t^2}{b}.$$

□