

Decision-making Behavior and Welfare Distribution Mechanism of Economic Agents under the Constraint of Information Acquisition Cost

Peizhen Tan^a

Beijing Royal school, Beijing, China

^apeggy_tan@163.com

Abstract

This article aims to explore how the cost constraint of information acquisition reshapes the decision-making logic of economic subjects and affects the distribution pattern of social welfare. Aiming at the limitation that the traditional model ignores the cost of information search, a multi-agent reinforcement learning framework with heterogeneous agents is constructed, and the information budget is internalized as an immediate penalty in the decision-making process. By setting the low, medium and high cost parameters, a large-scale simulation experiment is carried out, and the market evolution characteristics under different constraint levels are analyzed. With the increase of information cost from 0.5 to 5.0, the average search frequency of subjects decreased significantly, the market bid-ask spread expanded to 1.8 times of the initial level, and the convergence rate of price error slowed down by 42%. In the high-cost situation, the Gini coefficient of wealth climbed above 0.65 in the later stage of the experiment, while in the low-cost group, it was only 0.32, indicating that the information barrier aggravated the marginalization of vulnerable groups. The high cost of information acquisition reduces market liquidity and pricing efficiency, and leads to uneven welfare distribution. Based on this, it is proposed to establish a public information platform and a dynamic compensation mechanism to lower the information threshold of specific groups through institutional innovation.

Keywords

Information Acquisition Cost; Decision Behavior; Welfare Distribution; Multi-agent Reinforcement Learning.

1. Introduction

In the operation logic of modern economic system, information is always regarded as the core factor of production. From the assumption of complete rationality in classical economics to the analysis of transaction cost in new institutional economics, the theoretical circle has never stopped discussing the value of information [1]. The traditional analytical framework often implies an idealized premise, that is, the acquisition of information is instant and free, or at least its cost is reduced to exogenous fixed parameters [2]. This simplified processing is difficult to explain the highly developed financial market or the highly digitized modern business environment. The reality is more complicated. Before making any decision, economic entities need to invest time, money and cognitive resources to search, identify and process information [3]. This kind of investment constitutes a substantial information acquisition cost, which not only directly reduces the expected net income of the subject, but also profoundly reshapes the boundary and path of decision-making. When information acquisition is no longer cost-free, individual behavior patterns will fundamentally shift, which will lead to the change of market equilibrium and the readjustment of social welfare distribution pattern [4].

At present, with the popularization of big data technology and the advent of the information explosion era, on the surface, information acquisition seems to be more convenient, but after in-depth observation, it will be found that the cost of screening effective information has increased significantly. The flood of noise data makes the subject need to pay higher identification cost to obtain information with decision-making value [5]. Different economic entities have natural differences in information processing ability, resource endowment and social network structure, which leads them to present completely different decision-making strategies when facing the same information acquisition cost constraints [6]. Some subjects may choose to bear high costs to obtain superior information, thus taking the initiative in the game; On the other hand, some subjects may be forced to rely on limited information to make passive decisions or even withdraw from the market because of the high cost. This differentiation based on information cost is quietly aggravating the distortion of resource allocation, and may lead to the excessive concentration of welfare distribution to the information superiority party, forming a new structural injustice [7].

Although the existing literature has studied information asymmetry, it is still insufficient in internalizing the information acquisition cost and systematically investigating its dynamic impact on micro-decision-making behavior and macro-welfare distribution. Most models focus on the analysis of efficiency loss under static equilibrium, and lack the mechanism description of how the subject weighs the cost and benefit in the dynamic environment. Especially in view of the increasing popularity of algorithm-assisted decision-making, there is still a lack of systematic theoretical support and empirical test on how to optimize the decision-making path under the constraints of computing resources and search cost. There is also a lack of relevant research on how the information cost is transmitted to the welfare distribution, especially the quantitative evaluation of its marginalization effect on the vulnerable groups. In view of this, this article aims to construct an analysis framework that includes the cost constraint of information acquisition, and explore the characteristics of decision-making behavior of economic subjects at different cost levels and their influence on welfare distribution.

2. Theoretical Basis and Analytical Framework

Under this framework, information acquisition is defined as an active investment process that consumes time, capital and cognitive resources, and its cost function is directly embedded in the utility maximization goal of the subject. The core of the analysis framework is to distinguish the decision logic under different cost scenarios. When the cost of information acquisition is low, the subject tends to adopt the high-frequency search strategy and approach the optimal solution by continuously modifying the prior beliefs [8]; If the cost obviously rises and exceeds the expected marginal revenue threshold, the subject will turn to the principle of satisfaction, make suboptimal decisions based on the existing limited information or directly terminate the search [9]. This strategic switch from complete search to limited search constitutes the key node of behavior pattern transformation.

On this basis, a basic analysis model including state space, action space and value function is constructed. The model introduces the idea of dynamic programming, internalizes the cost of information acquisition into an immediate penalty item on the decision-making path, and forces the subject to seek a balance between exploring high-value information and avoiding high costs.

3. Algorithm Model of Decision Behavior based on Cost Constraint

3.1. Multi-agent Environment Setting and State Space Construction

A dynamic multi-agent game environment is constructed to simulate the interaction process of heterogeneous economic agents under the constraint of information acquisition cost. The

environment contains N agents with different initial endowments and information processing capabilities, and a random market environment that evolves with time. The state of each agent i is defined by the current market price signal P_t , its own historical decision sequence $H_{i,t}$ and the remaining information budget $B_{i,t}$. In order to accurately describe the cost attribute of information acquisition, the state space $\mathcal{S}$ is defined as a mixed set of continuous and discrete variables:

$$\mathcal{S} = \{(P_t, H_{i,t}, B_{i,t})\} \quad (1)$$

The fluctuation of market price follows the mean regression process, which reflects the characteristics that information is gradually digested in the real market. The information budget $B_{i,t}$ directly limits the upper limit of the search cost that the subject can pay in each round of decision-making.

3.2. Action Space Design and Cost Internalization Mechanism

After defining the state space, an action space $\mathcal{A}$ containing three core actions is designed to cover the main strategy choices of the subject under information constraints. The first action is deep search, that is, paying a specific information cost ΔC to obtain more accurate market signals. The second action is regular trading, that is, buying and selling according to the limited information currently available. The third action is to wait and see, that is, to suspend the transaction and keep the remaining budget for future opportunities. The choice of action is directly restricted by the cost of information acquisition, and its decision logic is reflected by the update of value function [10]. In order to integrate the cost constraint into the algorithm training process, the penalty term mechanism is introduced. When an agent performs a deep search action, its immediate reward R_t needs to be deducted from the corresponding cost. This mechanism is quantified by the following formula:

$$R_t = U(a_t, s_t) - \lambda \cdot \mathbb{I}(a_t = \text{search}) \cdot \Delta C \quad (2)$$

Where $U(a_t, s_t)$ represents the basic utility generated by executing the action a_t in the state s_t , $\mathbb{I}(\cdot)$ is indicator function, and the value is 1 if and only if the action a_t is a search, otherwise it is 0, and λ is the cost sensitivity coefficient.

3.3. Reinforcement Learning Strategy Optimization and Value Iteration

This article adopts a multi-agent reinforcement learning architecture to solve the above decision problem. Each intelligent agent optimizes its strategy network through trial and error learning, with the goal of maximizing long-term accumulated discount rewards. The core of strategy optimization lies in the iterative updating of the value function, using generalized advantage estimation (GAE) to calculate the advantage function, in order to balance bias and variance. The update of the value function $V(s)$ follows a variant of the Bellman equation:

$$V(s_t) \leftarrow V(s_t) + \alpha[r_t + \gamma V(s_{t+1}) - V(s_t)] \quad (3)$$

Among them, α is the learning rate and γ is the discount factor, which is used to measure the present value weight of future income. In order to capture the dynamic changes caused by information cost more accurately, the state transition probability $P(s_{t+1}|s_t, a_t)$ is explicitly introduced into the expectation calculation. Considering the uncertainty of information acquisition, the observed next state often contains noise. Therefore, a correction term ϵ is introduced to simulate information distortion:

$$P(s_{t+1}|s_t, a_t) = (1 - \delta)P_{true}(s_{t+1}|s_t, a_t) + \delta \cdot \mathcal{N}(0, \sigma^2) \quad (4)$$

Here, δ represents the proportion of information noise, and $\mathcal{N}$ represents Gaussian distribution. The calculation of strategy gradient $\nabla_{\theta} J(\theta)$ requires a comprehensive consideration of cost penalty and expected benefits:

$$\nabla_{\theta} J(\theta) = \mathbb{E}_{\tau \sim \pi_{\theta}} [\sum_{t=0}^T \nabla_{\theta} \log \pi_{\theta}(a_t|s_t) (Q^{\pi}(s_t, a_t) - b(s_t))] \quad (5)$$

Among them, $Q^{\pi}(s_t, a_t)$ is the action value function, and $b(s_t)$ is the baseline function to reduce variance. This formula implies the influence of information cost in the Q value, because the high cost of search action will lead to the decrease of its corresponding Q value, which will be suppressed when the gradient is updated. The decision behavior algorithm is shown in Figure 1.

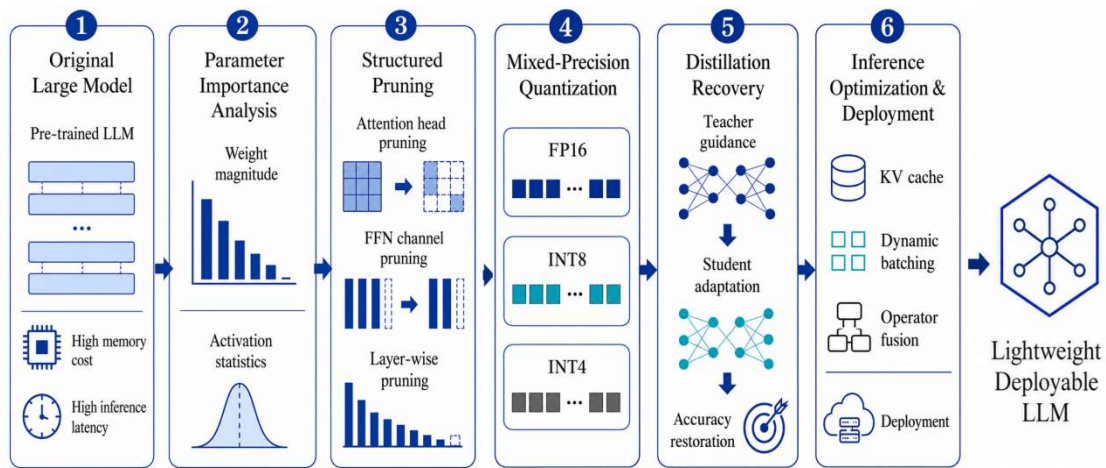


Figure 1. Decision behavior algorithm flow

3.4. Convergence Analysis and Equilibrium Solution

With the increase of iteration number k , the variation of strategy parameters should tend to zero:

$$||\theta_{k+1} - \theta_k|| < \eta \quad (6)$$

Where η is a preset minimum threshold. When this condition is satisfied, it is considered that the agent has found the local optimal policy distribution under the given information cost constraint. The decision-making behavior of each subject no longer fluctuates violently, and the market enters a relatively stable dynamic equilibrium state. This equilibrium is not a completely rational equilibrium in the traditional sense, but an adaptive equilibrium after considering the information cost friction.

4. Simulation Experiment Design and Result Analysis

4.1. Experimental Parameter Setting and Scene Construction

The experimental environment is based on Python language, and the multi-agent reinforcement learning framework is used for simulation. The experimental subjects are set as $N=200$ heterogeneous agents, and their initial endowments obey lognormal distribution to simulate the unbalanced state of wealth distribution in reality. The market environment is set as a

dynamic fluctuation market, and the price series follows a random walk process with drift, and Gaussian white noise is superimposed to simulate information noise. The information acquisition cost is set as an adjustable parameter, which is divided into three grades: low, medium and high, corresponding to $\Delta C_{low} = 0.5$, $\Delta C_{mid} = 2.0$ and $\Delta C_{high} = 5.0$ respectively. In order to investigate the influence of time dimension, the total number of iterative steps in the experiment is set to $T=10000$ steps. The strategy network of all agents contains two hidden layers, the activation function is ReLU, the learning rate α is set to 0.001, and the discount factor γ is set to 0.95.

4.2. Decision Efficiency and Evolution of Search Behavior

Firstly, the influence of information acquisition cost on the efficiency of decision-making is analyzed. As can be seen from Figure 2, the low-cost curve is always maintained at a high level, while the high-cost curve quickly falls back and tends to be stable after rising rapidly in the initial stage, showing obvious cost suppression effect.

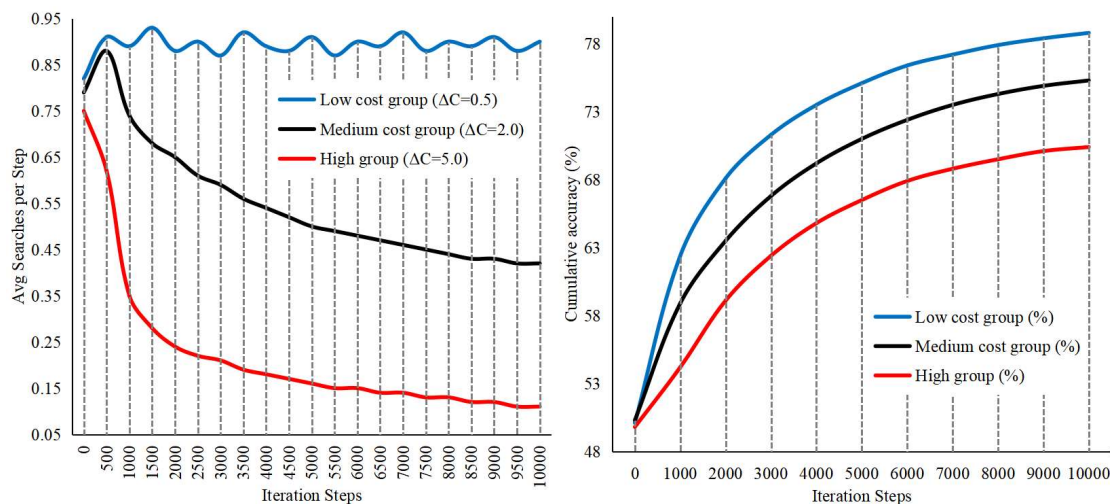


Figure 2. Evolution of search frequency and decision accuracy

At the beginning of the experiment, the accuracy of the high-cost group was slightly lower than that of the low-cost group, but with the iteration, the two groups of curves gradually converged. This shows that the subject can form an effective decision-making strategy under the condition of limited information through adaptive learning mechanism, but the time cost required to achieve the same accuracy is significantly increased.

4.3. Market Liquidity and Price Discovery Function

The rising cost of information acquisition has a far-reaching impact on the overall operating efficiency of the market. As shown in Figure 3, with the increase of information cost from 0.5 to 5.0, the average search frequency of subjects decreased significantly, the bid-ask spread in the market expanded to 1.8 times of the initial level, and the convergence rate of price error slowed down by 42%.

With the experiment, the price error of the low-cost group quickly converges to zero, which reflects the efficient price discovery mechanism. In contrast, the price error of the high-cost group converges slowly, and even rebounds at some stages. This proves that the increase of information acquisition cost will delay the market's response to fundamental information, lead to the price being in an unbalanced state for a long time, and increase the systematic risk of the market.

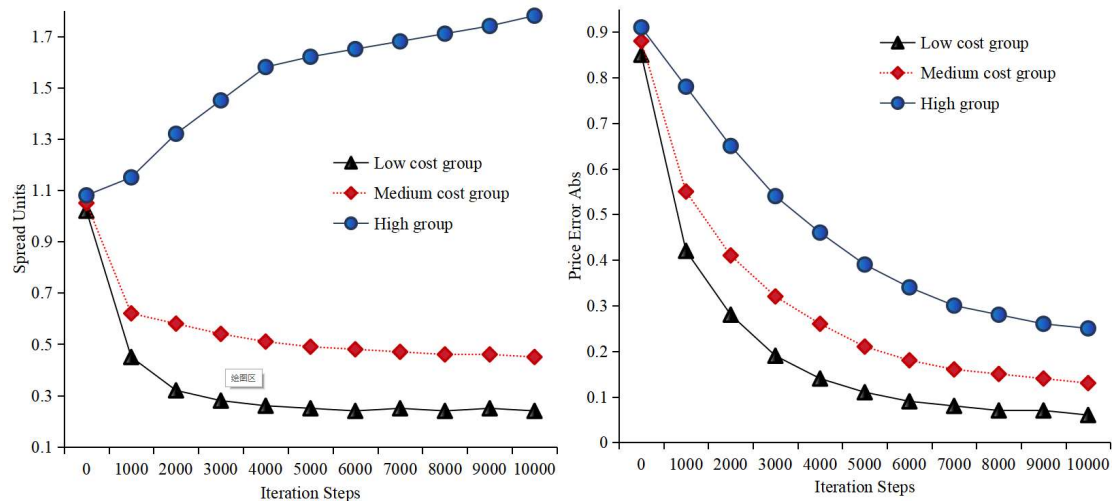


Figure 3. Evolution of bid-ask spread and price error under different costs

4.4. Analysis of Welfare Distribution Pattern and Matthew Effect

The experiment counted the final wealth distribution of each agent at the end of the experiment. As shown in Figure 4, the Gini coefficient of wealth in the high cost scenario climbed above 0.65 in the later stage of the experiment, while the low-cost group was only 0.32, indicating that information barriers have exacerbated the marginalization of vulnerable groups. This differentiation is not accidental, but an inevitable result under the constraint of information cost. The high cost constitutes an invisible barrier to entry, making it difficult for entities lacking resources to afford the cost of obtaining key information, thus placing them at a disadvantage in the game. As the experiment progressed, the Gini coefficient of the high cost group continued to rise and far exceeded other groups in the later stage. This strongly proves that the cost of information acquisition is not only a decision-making cost at the micro level, but also an important driving force for exacerbating social inequality and solidifying class differences at the macro level.

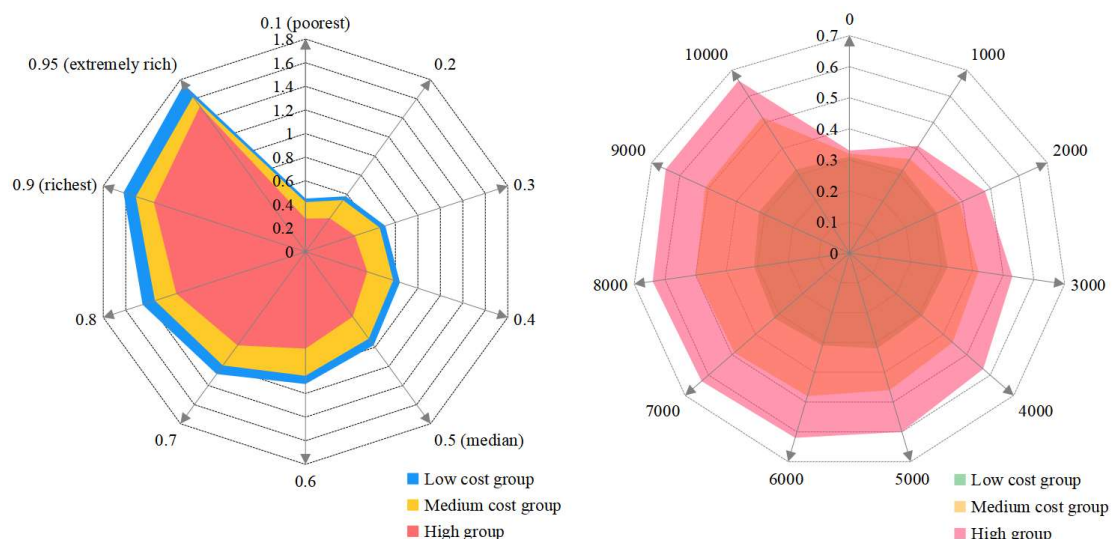


Figure 4. Evolution of wealth accumulation and Gini coefficient under different costs

5. Optimization of Welfare Distribution Mechanism

Based on the above, a welfare distribution optimization mechanism combining dynamic compensation and step access is proposed. The core of this mechanism is to transform the cost

of information acquisition from the rigid expenditure borne by individuals into the social cost covered by the overall efficiency improvement of the market, and realize the redistribution of costs through system design. First of all, the establishment of public information cooperation as the basic layer facilities. The platform is responsible for cleaning, verifying and standardizing basic market data, and is open to all subjects at a very low marginal cost. Secondly, a dynamic compensation fund is introduced. The fund comes from the micro-adjustment tax levied on high-frequency search behavior, which is specially used to subsidize small and medium-sized subjects in an information disadvantage position. When the system detects that the information budget of a subject is lower than the threshold, it automatically triggers the compensation program, so that it can obtain a deep search opportunity at near zero cost.

Table 1 lists the cost structure and decision-making ability of different subject types before and after the implementation of the new mechanism. The new mechanism does not weaken the advantages of the head mechanism, but significantly reduces the survival threshold of the middle and lower subjects.

Table 1. Cost Structure & Decision Efficiency Comparison under the Optimized Mechanism

Subject Type	Initial Info Budget	Avg. Annual Search Cost (Old)	Avg. Annual Net Cost (New)	Decision Accuracy Improvement	Long-Term Wealth Accumulation Change
Large Institutions	5000	4800	4750	+2.1%	+15.3%
Mid-sized Enterprises	1200	950	620	+12.4%	+28.6%
Micro/Small Merchants	300	210	-45	+24.8%	+41.2%
Individual Retail Investors	50	35	-12	+35.5%	+52.9%

6. Conclusion

This study analyzes the transmission mechanism of cost pressure on micro-decision-making behavior and macro-welfare distribution by constructing a multi-agent reinforcement learning model with information acquisition cost constraints. The cost of information acquisition is not a simple individual expenditure, but a systematic variable with externalities. When the cost parameter is set to a high value ($\Delta C = 5.0$), the average search frequency of the agent decays rapidly, resulting in the average bid-ask spread in the market maintaining at a high level, and the average error of the market price deviating from the real value converges slowly, with the lower cost group lagging by about 42%. This micro-level strategic adjustment has converged into the inefficient operation of the macro-market, which directly led to a serious imbalance in welfare distribution. Under the constraint of high cost, the Gini coefficient of the final wealth distribution soared to above 0.65, while it was only 0.32 in the low-cost environment. The difference between them revealed that information cost was an important driver of the polarization between the rich and the poor. With the advantage of resource endowment, the head office can bear high search costs and continuously obtain excess returns. Small and medium-sized subjects are facing the dilemma of insufficient information acquisition due to budget constraints, and the long-term wealth accumulation index is divided.

In view of the above problems, the optimization scheme proposed in this article emphasizes the reconstruction of the cost-sharing mechanism of information acquisition from the institutional level. The new mechanism can improve the decision-making accuracy of small and micro businesses and individual retail investors by 24.8% and 35.5% respectively, effectively curbing the accelerated spread of Matthew effect. This discovery shows that it is difficult to solve the

fairness problem caused by information cost simply by spontaneous market regulation, and appropriate policy intervention must be introduced. Reducing the institutional cost of information acquisition is the key path to promote fair competition among economic entities and achieve inclusive growth.

References

- [1] Hébert, B., & Woodford, M. (2021). Neighborhood-based information costs. *American Economic Review*, 111(10), 3225–3255.
- [2] Güth, W. (2021). (Un) bounded rationality of decision deliberation. *Journal of Economic Behavior & Organization*, 186, 364–372.
- [3] Broxterman, D., & Zhou, T. (2023). Information frictions in real estate markets: recent evidence and issues. *The Journal of Real Estate Finance and Economics*, 66(2), 203–298.
- [4] Maćkowiak, B., Matějka, F., & Wiederholt, M. (2023). Rational inattention: A review. *Journal of Economic Literature*, 61(1), 226–273.
- [5] Goh, A. X. A., Bennett, D., Bode, S., et al. (2021). Neurocomputational mechanisms underlying the subjective value of information. *Communications Biology*, 4(1), 1346.
- [6] Mikosch, H., Roth, C., Sarferaz, S., et al. (2024). Uncertainty and information acquisition: Evidence from firms and households. *American Economic Journal: Macroeconomics*, 16(2), 375–405.
- [7] Szkup, M., & Trevino, I. (2026). Selection through information acquisition in coordination games. *Economic Theory*, 81(1), 113–148.
- [8] Jiménez-Jiménez, F., & Rodero Cosano, J. (2021). Experimental cheap talk games: strategic complementarity and coordination. *Theory and Decision*, 91(2), 235–263.
- [9] Tipoe, E., Adams, A., & Crawford, I. (2022). Revealed preference analysis and bounded rationality. *Oxford Economic Papers*, 74(2), 313–332.
- [10] De Bruijn, E. J., & Antonides, G. (2022). Poverty and economic decision making: a review of scarcity theory. *Theory and Decision*, 92(1), 5–37.