Estimating information cost functions in models of rational inattention

By Ambuj Dewan and Nathaniel Neligh

Introduction

- In recent literature it became more prominent that people's limited capacity to pay attention plays an important role in decision making
- While this may be evidence for bounded rationality, it can also be that dealing with information is costly (gathering, synthesizing, thinking takes effort and time), thus a decision-maker not engaging in this activity could be a result of a rational decision.
- Models dealing with costly information where individuals rationally do not pay attention are called *Rational inattention* models (Sims 2003)
- In these models:
    - 1st decision makers (DM) choose which information to acquire, and they pay the costs
    - 2nd They evaluate their choices based on the information they possess, then they make the decision.
- Rational Inattention models are widely used, but authors make different assumptions in their models regarding the cost function of acquiring information
- This paper aims to test the DM's information costs functions in a lab setting, and evaluate which assumptions hold for these functions. It also evaluates widely used functional forms in Rational Inattention literature.

Watered down theoretical framework (For detailed derivations see the paper, (and appendix))

- The world has multiple states (n unique states), and each state has equal probability
- The DM tries to guess the state of the world, let qi,j denote the probability that DM guesses the state of the world is i, while the actual state of the world is j. (I and j can be 1…n)
- The performance (P) of the DM can be shown as:

$$P = \frac{1}{n}\sum_{i=1}^{n} q_{i,i}$$

  That is Performance (P) is the probability of guessing correctly averaged across all states of the world
- DM maximizes r*P-C(P), by choosing her performance, where r is the reward and C is the cost, that is performance dependent. All in all, DM chooses costly performance to maximize her payoffs.
- This choice P(r) is the performance function which maps from incentives (rewards) to performance.
- DM is rationally inattentive iff "no improving attention cycles" (NIAC) and "no improving action switches" (NIAS) conditions are satisfied.
- **Proposition 1**. *The DM's behavior is consistent with NIAC iff P(r)is non-decreasing in r.*
- If rewards are higher DM does not decrease effort.
- **Proposition 2.** *The DM's behavior is consistent with NIAS iff for all incentives, if agent guessed the state of the world to be x, then no other state of the world is more likely than x.*
- That is there is no reason for DM to change her guessing strategy.

- *Definition 1. A DM is responsive (to incentives) in a uniform guess task if for some r2>r1, P(r2) >P(r1).*
- That is DM does not perform the same way at all incentive levels.
- **Cost functions**
- If the cost function is well-behaved (strictly convex, continuous, strictly increasing) and differentiable, the increasing part of the performance function is equal to the inverse of the derivative of the cost function.
- Based on this, if we know the performance function, we can get back the cost function
- Commonly used cost functions:

1. ***Quadratic cost function***
    - Main reason: implies linear performance function up to maximum performance

$$C(P) = \begin{cases} 0, & P \leq d \\ c(P-d)^2, & P > d \end{cases}$$

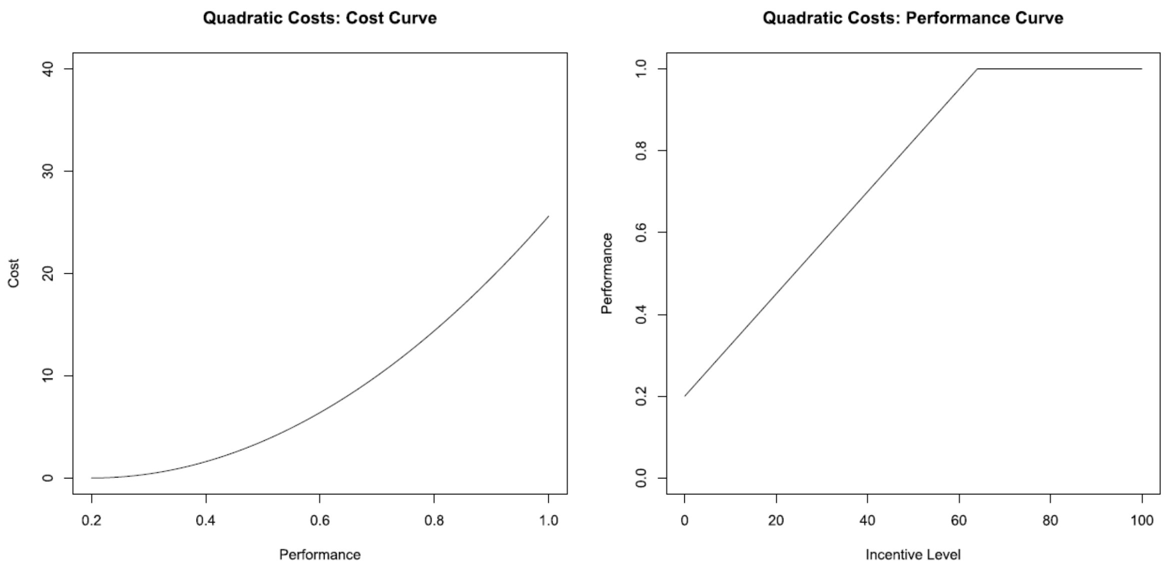


Fig. 1. Quadratic costs. The left panel shows the cost function for $c = 40$ and $d = 0.2$, and the right panel shows the resulting performance curves.

2. ***Entropy based cost functions***
    - Main reason: Entropy is a measure of information content/uncertainty of a random variable, so it follows (at first) naturally that costs are associated with information content.
    - Shannon entropy:

$$H^S(p) := -\alpha \sum_{i=1}^{n} p_i \ln(p_i)$$

-
- Tsallis entropy (generalized entropy) if sigma -> 1 we get back Shannon entropy:

$$H^T(p) := \frac{\alpha}{\sigma - 1}\left(1 - \sum_{i=1}^{n} p_i^{\sigma}\right)$$

-

- Costs:

$$C(P)=\begin{cases}\dfrac{\alpha}{\sigma-1}\left[P^{\sigma}+(n-1)^{1-\sigma}(1-P)^{\sigma}-n^{1-\sigma}\right], & \sigma\neq 1\\ \alpha\left[\ln(n)+P\ln(P)+(1-P)\ln\left(\dfrac{1-P}{n-1}\right)\right], & \sigma=1\end{cases}$$

-

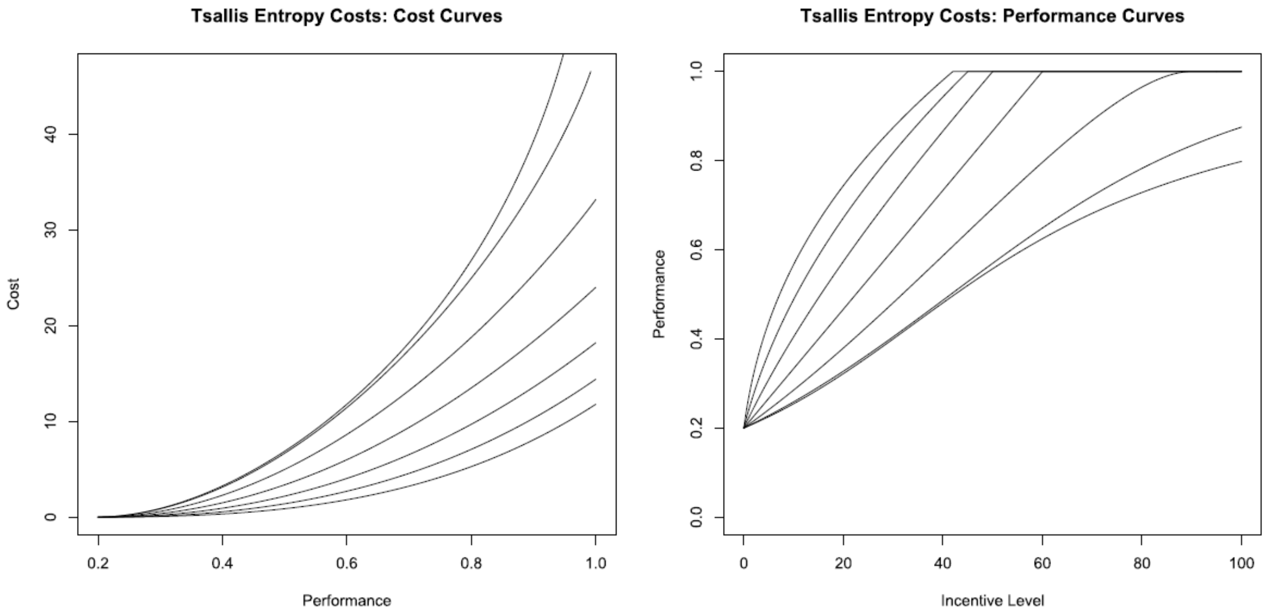


Fig. 2. Tsallis entropy costs. The left panel shows the cost function for $\sigma \in \{0.5, 1, 1.5, 2, 2.5, 3, 3.5\}$, going clockwise, and the right panel shows the resulting performance curves for those values of σ, going counterclockwise. α is set at 30.

3. *Normal signals*

- Main reason: DM receives normally distributed signals of the true state of the world, costly effort reduces the variance of these signals.
- Ugly formula

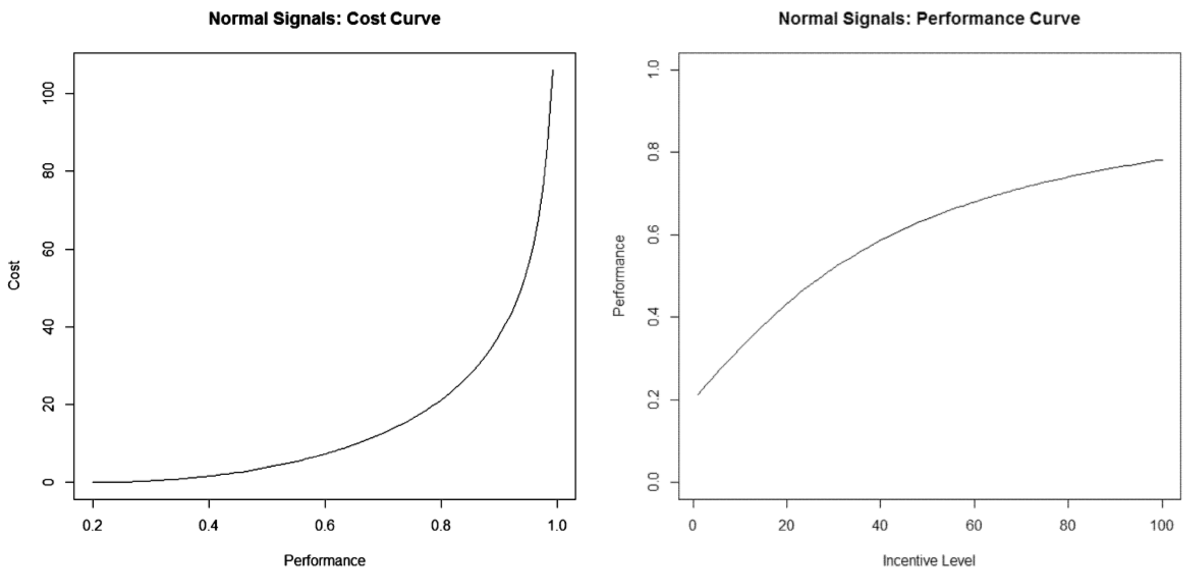


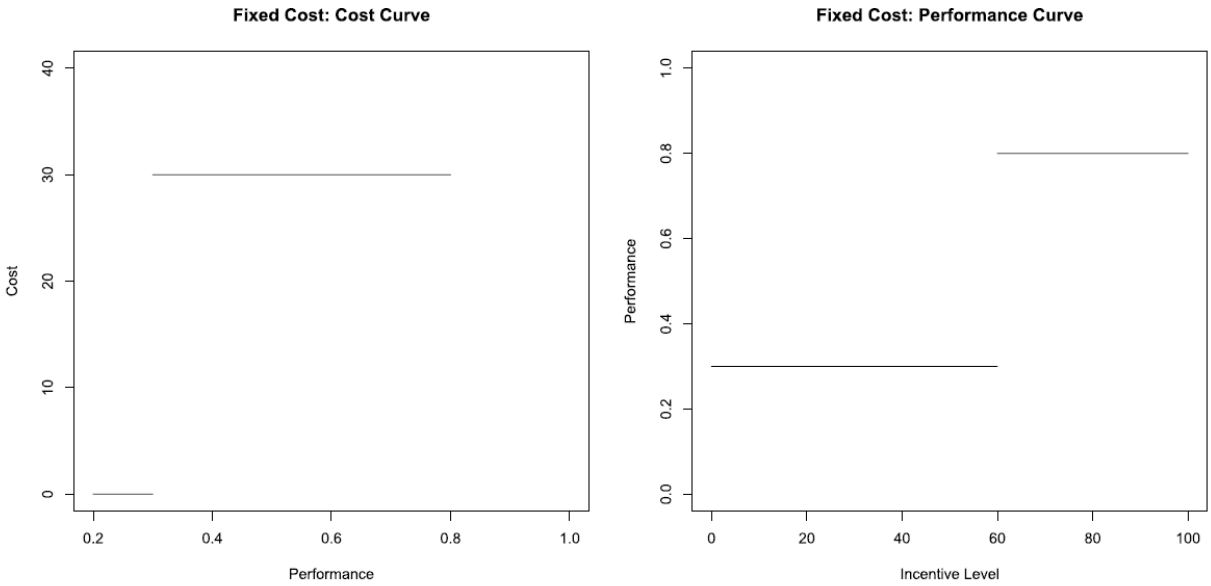
Fig. 3. Normal signals with cost of precision given by $K(\zeta) = 4\zeta^2$. The left panel shows the cost function, and the right panel shows the resulting performance curve.

4. *Fixed costs*

- Main idea: some information for free have to pay a fixed cost to receive more information

$$C(P) = \begin{cases} 0, & P \leq \underline{q} \\ \kappa, & P \in (\underline{q}, \bar{q}] \\ \infty, & P > \bar{q} \end{cases}$$

-



5. *Hybrid costs*

- Main idea: free initially, not costly up to a point, more effortful up to another point. (Thinking passively, and actively)
-

$$C(P) = \begin{cases} 0, & P \leq d_1 \\ c_1(P - d_1)^2, & d_1 < P \leq d_2 \\ c_1(d_2 - d_1)^2, & d_2 < P \leq d_3 \\ c_2(P - d_3)^2 + c_1(d_2 - d_1)^2, & P > d_3 \end{cases}$$

-

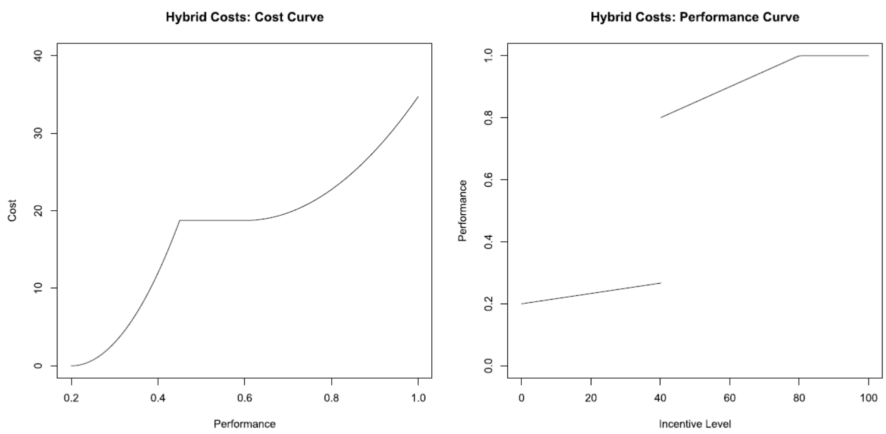


Fig. 6. Hybrid costs. The left panel shows the cost function, and the right panel shows the resulting performance curve. Parameters are $c_1 = 300$, $c_2 = 100$, $d_1 = 0.2$, $d_2 = 0.45$, and $d_3 = 0.6$.

- Summary of cost functions:

Table 2
Properties of cost functions.

Cost Function	Continuous	Convex	Performance Function
Differentiable and well-behaved	Yes	Yes	Inverse of derivative
Quadratic	*Yes*	*Yes*	*Affine*
Tsallis entropy	Yes	Yes	Sigmoid/inverse-S/concave (SIC)
Mutual information	*Yes*	*Yes*	*Logistic*
Normal signals	Yes	Yes	Concave
Dual-process	Can be	No	Discontinuous
Fixed costs	*No*	*No*	*Binary*
Hybrid	*Yes*	*No*	*Piecewise affine*

Note: Performance-function properties of normal-signal costs are for state spaces with equidistant spacing.

Experimental design

- Series of perceptual tasks
- Random arrangement of dots, 38-42 each equally probability.
- DM has to choose between options 38-42 dots.
- Reward is points: 83 points means 0.83 chance of getting a reward at the experiment.
- Two rewards were realized at the end of the experiment.

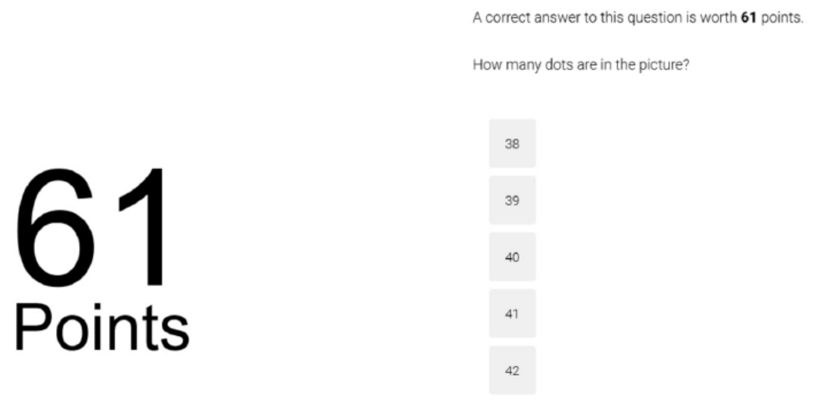


Fig. 7. Incentive display for a task.

- Incentive levels first increasing from 1 to 99 then by 2, then decreasing from 100 to 2 by 2.

- Experiment is designed, so multiple options are present (can be a measure of how bad did DM guess)
- No gambles, makes everything more simple
- Reward structured in a way, that does not matter for risk preference

Results

- 81 subject overall
- Subset structure: All subjects > Rationally inattentive > Responsive > Well-behaved
- Problem: each incentive level measured once performance is problematic (either 1 or 0), they solve this by using bins of incentive levels (10 wide)
- Rationally inattentive:
- reject NIAC if quadratic polynomial fit is not monotonic increasing (5% confidence)
- 4 rejected
- Reject NIAS using a bootstrap procedure (5% confidence, they test a weaker assumption than actual NIAS)
- 7 rejected
- Overall, 70 out of 81 is considered rationally inattentive (fail to reject both NIAS and NIAC)
- Responsiveness:
- Reject responsiveness if performance is constant (coefficient of incentive level is not significantly different from 0 at the 5% level) in all incentive levels, incentive levels 1-50 and incentive levels 51-100
- 42 out of 70 categorized as responsive.
- Well-behaved
- Likelihood ratio test between a step function and a logistic model.
- 29 out of 42 may have discontinuity, for 13 discontinuity is rejected

Table 3
Categorization of subjects.

Category	Of All Subjects	Of R.I. Subjects	Of Resp. Subjects
All subjects	81 (100%)	—	—
R.I. subjects	70 (86.4%)	70 (100%)	—
Resp. subjects	///	42 (60.0%)	42 (100%)
W.B. subjects	///	///	13 (31.0%)

Note: "R.I." = rationally inattentive; "Resp." = responsive; "W.B." = well-behaved, i.e. subjects whose behavior is consistent with continuous, convex cost functions. — denotes that the column category is a subset of the row category, and /// denotes that the row category is defined only on a subset of the column category.

-

Model selection

- Each model is estimated with a different method, Akaike Information Criteria (AIC) is compared (lower the score the better the model is)

Table 6
Model Selection for Responsive Subjects.

Model	Binary (2)	Logistic (7)	SIC (8)	Concave (9)
Number of Subjects	10 (23.8%)	19 (45.2%)	7 (16.7%)	6 (14.3%)

-

Table 7
Average AIC and Rank for Estimated Models.

	Model	AIC	Rank
1	Constant	131.165	7.476
2	Binary	114.060	2.881
3	Affine with break	119.982	4.429
4	Affine	117.832	5.833
5	Quadratic	131.123	6.595
6	Cubic	132.929	7.643
7	Logistic	116.008	2.714
8	SIC	113.672	2.310
9	Concave	121.967	5.119

-
- Binary and SIC seem to be the best fit.

Conclusion

- Most subjects are rationally inattentive
- Big heterogeneity in behavior
- Tsallis entropy seems to the best if one wants to model costs with a single model
- Both binary and continuous performance functions are present