Attention Constraints and Learning in Categories

Peiran Jiao [1] Rahul Bhui [2]

[1]Maastricht University & Nuffield College

[2]Massachusetts Institute of Technology

EEA-ESEM 2022
Milan

Motivation: Limited Attention

- Attention is a scarce cognitive resource.

Motivation: Limited Attention

- Attention is a scarce cognitive resource.
- Examples: investors(e.g. Huberman and Regev, 2001; Hirshleifer and Teoh, 2003), consumers (Gabaix and Laibson, 2003; Gabaix et al., 2003), firms (Hirshleifer et al., 2003).

Motivation: Limited Attention

- Attention is a scarce cognitive resource.
- Examples: investors(e.g. Huberman and Regev, 2001; Hirshleifer and Teoh, 2003), consumers (Gabaix and Laibson, 2003; Gabaix et al., 2003), firms (Hirshleifer et al., 2003).
- Rational inattention? Or biased?

Motivation: Limited Attention

- Attention is a scarce cognitive resource.
- Examples: investors(e.g. Huberman and Regev, 2001; Hirshleifer and Teoh, 2003), consumers (Gabaix and Laibson, 2003; Gabaix et al., 2003), firms (Hirshleifer et al., 2003).
- Rational inattention? Or biased?
- We ask: Do people *adjust* attention allocation in response to changes in the environment *optimally*?

Motivation: Categorization

- Limited attention $\Rightarrow$ Focus on a subset of information

Motivation: Categorization

- Limited attention $\Rightarrow$ Focus on a subset of information
- Which subset?
  - Salient information (e.g. Bordalo et al., 2012, 2013)
  - Easier-to-process information
    Such as category-level information (Our focus)

Motivation: Categorization

- Limited attention ⇒ Focus on a subset of information
- Which subset?
  - Salient information (e.g. Bordalo et al., 2012, 2013)
  - Easier-to-process information
    Such as category-level information (Our focus)
- Psychology: people categorize to facilitate predictions (e.g. Murphy, 2002).
- Tradeoff:
  - Coarser categories
  - Finer categories
- Example: color concepts

Implications

- Style investing (Barberis and Shleifer, 2003).
- Stock price comovement (Peng and Xiong, 2006).
- Superficial name change (Copper et al., 2001).

What we do

- A Model
- An experiment with
  - Predetermined categories
  - Processing time as a proxy for attention
  - Endogenous allocation

Model

- The representative investor holds a portfolio with m industries and n firms within each industry.
- Firm j in industry i pays a dividend $d_{j,i,t}$ in each period t.

$$d_{j,i,t} = h_t + f_{i,t} + g_{j,i,t} \tag{1}$$

- $i = 1, ..., m$ and $j = 1, ..., n$; h_t is the market component, $f_{i,t}$ is the industry component and $g_{j,i,t}$ is the firm-specific component.

Model

- These three components are independent from each other, and i.i.d. across periods.

$$
\begin{aligned}
h_t &\sim N(\bar{h}, \sigma_h^2), \\
f_{i,t} &\sim N(\bar{f}, \sigma_f^2),\ i = 1, ..., m, \\
g_{j,i.t} &\sim N(\bar{g}, \sigma_g^2),\ i = 1, ..., m,\ j = 1, ..., n.
\end{aligned}
$$

Model

- Exponential utility: $u(c) = -\frac{1}{\gamma}e^{-\gamma c}$
- An investor maximizes her lifetime expected utility

$$\max E_t[\sum_{s=t}^{\infty} \delta^{s-t} u(c_s)] \tag{2}$$

Model

- Exponential utility: $u(c) = -\frac{1}{\gamma}e^{-\gamma c}$
- An investor maximizes her lifetime expected utility

$$\max E_t[\sum_{s=t}^{\infty} \delta^{s-t} u(c_s)] \tag{2}$$

- Intertemporal Budget Constraint: Initial cash holdings plus current dividends are either consumed or carried over to next period.

$$K_{t+1} = (K_t - c_t)(1+r) + \sum_{i=1}^{m}\sum_{j=1}^{m} d_{i,j,t+1} \tag{3}$$

- Need to forecast dividends to better smooth consumption over time.

Model

- To forecast dividends: Need to process information.
- The investor reads press releases, media coverages, analyst reports, etc.,
- Attention Constraint: κ.

$$\lambda_{h,t} \in [0,1]$$
$$\lambda_{f,i,t} \in [0,1]$$
$$\lambda_{g,i,j,t} \in [0,1]$$

$$\lambda_{h,t} + \sum_{i=1}^{m} \lambda_{f,i,t} + \sum_{i=1}^{m}\sum_{j=1}^{n} \lambda_{g,i,j,t} \leq 1 \tag{4}$$

Model

- The investor processes information to extract signals, s for each factor.

$$
\begin{aligned}
s_{h,t} &= h_t + \epsilon_{h,t}, \quad \epsilon_{h,t} \sim N(0, \eta^2_{h,t}), \\
s_{f,i,t} &= f_{i,t} + \epsilon_{f,i,t}, \quad \epsilon_{f,i,t} \sim N(0, \eta^2_{f,i,t}), \\
s_{g,i,j,t} &= g_{j,i,t} + \epsilon_{g,i,j,t}, \quad \epsilon_{g,i,j,t} \sim N(0, \eta^2_{g,i,j,t}).
\end{aligned}
$$

- Errors are independently drawn.
- Need to capture: The more attention on a factor, the better signal on that factor.

Model

- Entropy method to measure informational value of signal.
  - The amount of information contained in s is measured by its entropy;
  - Entropy is the reduction of uncertainty due to s;
  - More attention to a factor $\Rightarrow$ Generate s with higher entropy.
- Tradeoff:
  - More attention to one factor
    $\Rightarrow$ More accurate signal on that factor;
    $\Rightarrow$ Better forecast on that factor;
    $\Rightarrow$ Less attention available for other factors.

Model Summary

- The investor maximizes lifetime expected utility.
- Given two constraints:
  - Attention Constraint
  - Budget Constraint
- Equilibrium:
  - Marginal contributions of attention to all factors are equalized.
  - Asset prices as the shadow prices from the pricing kernel are determined by marginal utilities.

Experimental Design: **Control**

A simplified prediction task

- There are only two levels: 1 industry and n firms.
- Each firm's dividend is a linear combination of the industry component and the firm's idiosyncratic component.

$$d_{j,t} = f_t + g_{j,t} \tag{5}$$

- f_t and $g_{j,t}$ follow normal distributions known to the subjects.

Experimental Design: Task

Each session contained 2 blocks, 10 periods per block.
Each period was divided into Observation and Prediction phases.

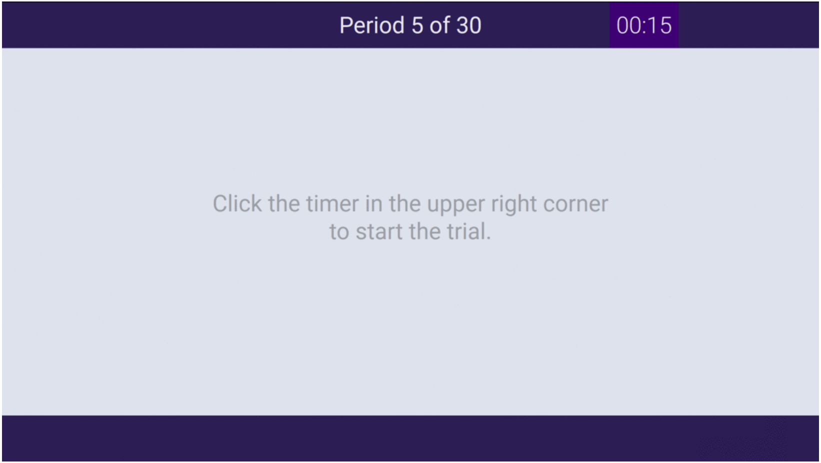
- Observation phase:
  - Mouse over: reveal signals of that component;
  - Move to another component: previous signals disappear;
  - Limited time: e.g. 15 seconds per period.
- Prediction phase:
  - Forecast next period values of all firms' dividends;
  - Accuracy rewards:
    Predictions of one randomly chosen period are rewarded;
    Calculated based on Mean Squared Errors.

What are the signals?

- When subjects move their mouse to a factor, a sequence of signals are revealed.

$$s_{f,t} = f_t + \epsilon_{f,t}, \quad \epsilon_{f,t} \sim N(0, \eta_\tau^2),$$
$$s_{g,j,t} = g_{j,t} + \epsilon_{g,j,t}, \quad \epsilon_{g,j,t} \sim N(0, \eta_\tau^2).$$

Screenshot


Period 5 of 30
00:15
Click the timer in the upper right corner
to start the trial.

Theoretical Predictions

- **Control**: Processing time
  - More time on the industry factor than on each individual firm
  - More time on a factor
    $\Rightarrow$ More accurate forecast on that factor
    $\Rightarrow$ Less accurate forecast on other factors
  - More time on a salient factor
  - Learning: time on industry stabilized but heterogeneous

Three Conditions

Within each condition, 1 parameter varies between blocks.

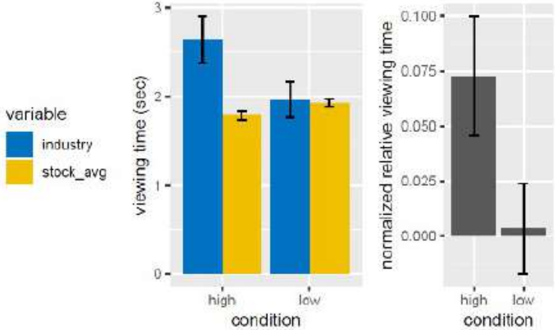
- **Variance**: High vs Low relative variance of Stocks.
  - More attention to industry factor under higher variance.
- **Size**: Many vs Few number of firms.
  - More attention to industry factor under more firms.
- **Time**: Lengthy vs Brief time each period.
  - More attention to industry factor under less time.

Procedure

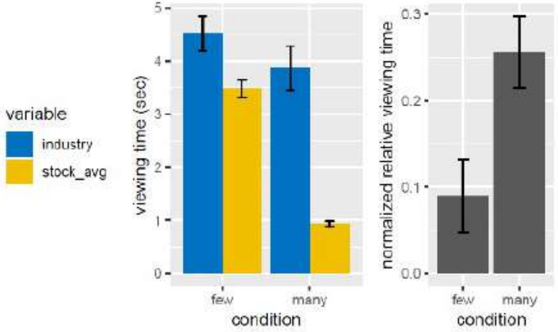
- Online experiment on Amazon Mechanical Turk.
  - Instructions and comprehension questions;
  - Main task with attention check;
  - Survey of demographics and financial literacy.
- A session took around 30 minutes.
- Subjects were paid $9 on average.

Results

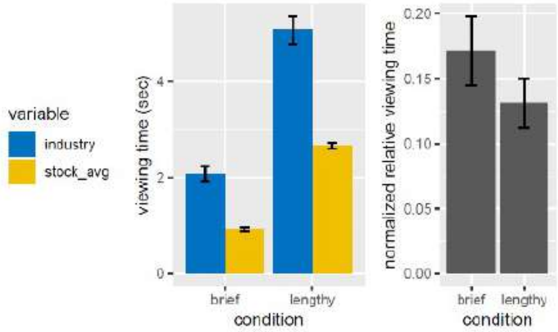
Attention Allocation: Variance



Attention Allocation: Size



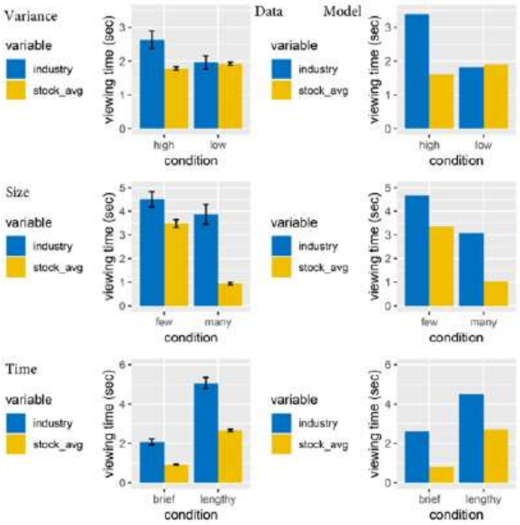
Attention Allocation: Time



Attention Allocation: Model

- 1 industry, n stocks
- Predict dividend of each stock i, which is $d_i = f + g_i$
- Priors: $f \sim N(0, \sigma_f^2)$, $g_i \sim N(0, \sigma_g^2)$; extract Gaussian signals
- Must spend attention budget κ to get information about factors
- Fraction α spent on industry, and $\frac{1-\alpha}{n}$ on each stock
- Entropy reduced by $I = \frac{1}{2}(\sigma^2_{prior}/\sigma^2_{posterior})$
- Suppose linear in attention, then $I = \frac{1}{2}\theta\kappa\alpha$
- Optimization problem: $\arg\min_\alpha \sum_i Var(V_i)$
- (Interior) solution: $\alpha^* = \frac{\frac{n}{\theta\kappa} log(\frac{n\sigma_f^2}{\sigma_g^2}+1)}{n+1}$

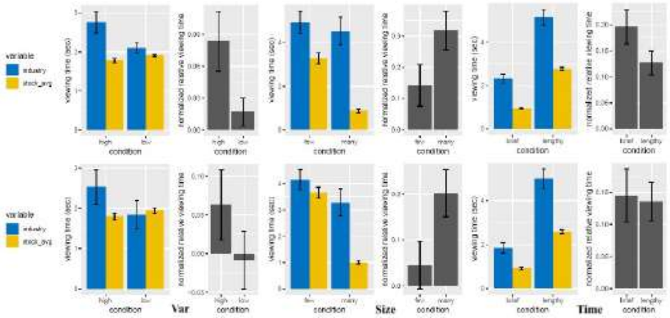
Attention Allocation: Model



Performance?

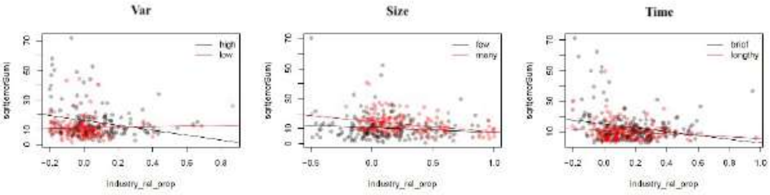
Accuracy

- Upper panels for High performers, and lower panels for Low performers.



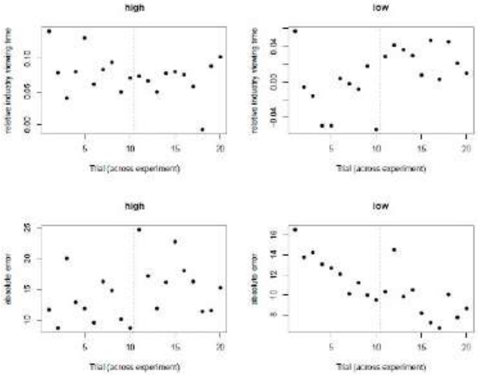
Accuracy and Attention Allocation

- Significant association between accuracy and (relative) attention to industry, but only when category is most important.

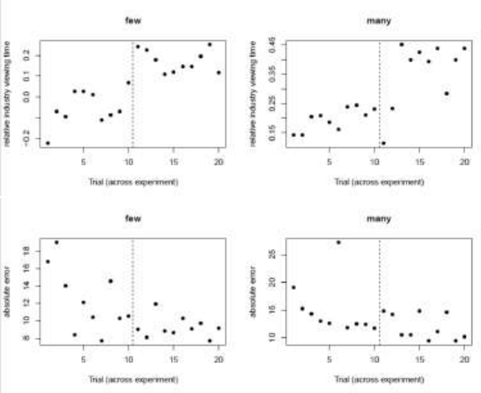


Learning?

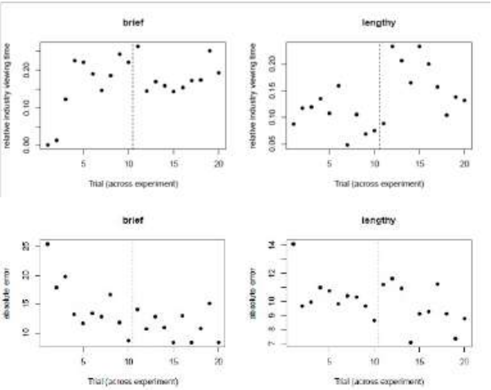
Learning: Variance



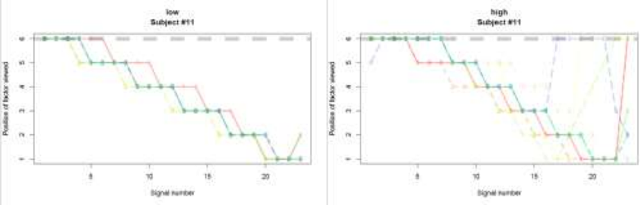
Learning: Size



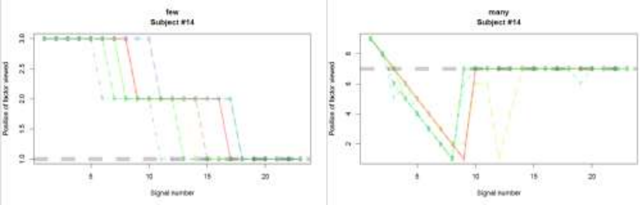
Learning: Time



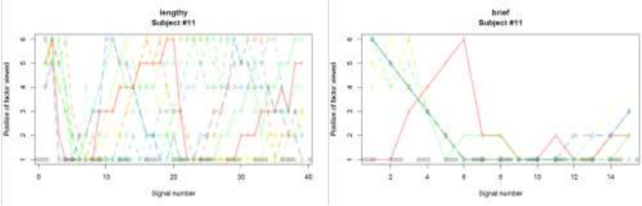
Attention Trajectory: Variance



Attention Trajectory: Size



Attention Trajectory: Time



Conclusion

- Attention shifts as predicted by model of optimal attention allocation
- Attention is more diffuse than predicted, but there are signs of convergence
- Some evidence of learning toward optimal behavior
- Future directions: Model of learning? Working memory? Overconfidence? Analysis of attention shift?

Thank you!
Questions Please.

Me:
http://peiranjiao.wordpress.com
p.jiao@maastrichtuniversity.nl
Link to paper:
https://doi.org/10.31234/osf.io/t5sy4