Introduction

- A stereotype is an “over-simplified image of [...] a type of person or thing” Oxford Dictionary

- Stereotypes are ubiquitous:
  - Racial groups (“Asian-Americans are good at math”)
  - Demographic groups (“Florida residents are elderly”)
  - Other groups (“ABS are safe”)

- Allow for quick assessments, but also distort judgment:
  - some are broadly accurate (dutch are tall), others inaccurate (irish red haired)
  - can lead to sub-optimal decisions
Several views on stereotypes

- **Statistical Discrimination** Phelps 1972, Arrow 1973
  - Rational expectation of a group’s types
  - Stereotypes are *accurate*

- **Sociology**
  - Focuses on pejorative and *inaccurate* social stereotypes
  - Often directed at minorities, and/or with economic motives
    Steele 2010, Glaeser, 2005

- **Social Psychology**
  - Stereotypes are “selective, […] localized around group features that are the most distinctive, that provide the greatest differentiation between groups, and that show the least within-group variation” Hilton and von Hippel 1996
This Paper

Model of stereotypes based on *representativeness*  
(Kahneman and Tversky 1972, 1974, 1983, Gennaioli and Shleifer 2010)

- “an attribute is representative of a class if it is very diagnostic; that is, the relative frequency of this attribute is much higher in that class than in a relevant reference class.”  
  
- Decision maker represents distribution $Pr(t|G)$ of types $t$ in $G$

- Most representative types come quickly to mind: 
  \[ \argmax_t \frac{Pr(t|G)}{Pr(t|-G)} \]
Example

- Age distribution of population of Florida vs US:

<table>
<thead>
<tr>
<th>age</th>
<th>0 – 18</th>
<th>19 – 44</th>
<th>45 – 64</th>
<th>65+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Florida</td>
<td>23.9%</td>
<td>31.6%</td>
<td>27.0%</td>
<td>17.3%</td>
</tr>
<tr>
<td>US</td>
<td>26.6%</td>
<td>33.4%</td>
<td>26.5%</td>
<td>13.5%</td>
</tr>
</tbody>
</table>

- Stereotype for Florida residents: \( t = 65+ \)
  - Maximizes \( \frac{\Pr(t|\text{Florida})}{\Pr(t|\text{US})} \)
  - but does not maximize \( \Pr(t|\text{Florida}) \): inaccurate stereotype

- In both populations, 19 – 44 is the most likely age bracket.
  - Florida residents have heavier elderly tail
  - Thus the elderly are the distinctive feature of Florida
Plan

For DMs whose quick thinking is governed by representativeness:

- characterise when stereotypes are accurate / inaccurate
  - “Texans are Republican” vs “Florida residents are elderly”
  - Some experimental evidence (abstract groups)

- characterise how stereotypes change
  - Neglect of data vs over-reaction to data

- Finance Application: stereotypical thinking and neglected risks
Setup:

- $G \subset \Omega$ is a group in the overall population $\Omega$
  - E.g. $G =$ Florida residents, $\Omega =$ Americans

- $T$ is a random variable – the “types” – defined over $\Omega$
  - $T$ is ordered, e.g. age, income, payoffs
  - categorical (later, continuous)

- DM forecasts conditional distribution $\left( \pi_{t,G} \right)_{t \in T}$ of $T$ in $G$
  - $\left( \pi_{t,G} \right)_{t \in T}$ is available in DM’s memory
  - $\left( \pi_{t,G}^{st} \right)_{t \in T}$ is stereotype of $G$

- Focus on close-ended questions where $T$ (and $\Omega$) are specified
The representativeness of type $t$ for group $G$ is

$$R(t, G) = \frac{\pi_{t,G}}{\pi_{t,-G}}, \quad \text{where} \quad -G = \Omega \setminus G.$$ 

Stereotype of $G$ is a truncation of $(\pi_{t,G})_{t \in T}$ to the $d$ most representative types (later see smoother discounting).

Let $T = \{t_1, \ldots, t_N\}$ be ranked by decreasing representativeness for $G$, and let $d \in \{1, \ldots, N\}$.

$$\pi_{t_r,G}^{st} = \begin{cases} \frac{\pi_{t_r,G}}{\sum_{r'=1}^{d} \pi_{t_{r'},G}}, & \text{for } r \in \{1, \ldots, d\}. \\ 0, & \text{otherwise.} \end{cases}$$
Representativeness vs Likelihood

- Representativeness $R(t, G) = \frac{\pi_{t,G}}{\pi_{t,-G}}$ is high when:
  - $t$ is more likely in $G$, i.e. $\pi_{t,G} - \pi_{t,-G}$ is large
  - $t$ is unlikely in comparison group, i.e. $\pi_{t,-G}$ is small
  - Stereotypes highlight (proportional) differences across groups

- Stereotype accuracy:
  - measured by a distance between distributions, such as:
    
    $$L[(\pi_{t,G}^{st})_{t=1,...,N}|(\pi_{t,G})_{t=1,...,N}] = \sum_{t}(\pi_{t,G}^{st} - \pi_{t,G})^2$$
  - accurate if likely types are representative, i.e. unlikely for $-G$

- Intuition: Stereotypes are accurate when groups are different, inaccurate when groups are similar
Stereotype Accuracy

Example 1. Voting in Hawaii and in Texas

Voting patterns in 2012 Presidential Election

- Democrat is the most representative and likely vote in Hawaii.
- Republican is the most representative and likely vote in Texas.
- Stereotypes are accurate:
  - For each group, representativeness increases with likelihood.
  - Groups are different (probability mass in different types)
Stereotype Accuracy

Example 2. Income distribution in the US

- Top income bracket is representative of Whites, but unlikely.
- Bottom bracket is representative of Blacks, but is not modal.
- Stereotypes are inaccurate:
  - For each group, representativeness differs from likelihood
  - Groups are similar (large probability mass in same types)
Extreme Stereotypes

- Suppose types are ordered $t_1 < \ldots < t_N$
  and $\pi_{t,G}/\pi_{t,-G}$ is monotonic in $t$ (MLRP holds).
  - if ratio is increasing, the stereotype is $\{t_{N-d+1}, \ldots, t_N\}$, and
    \[ \mathbb{E}^{st}(t|G) > \mathbb{E}(t|G) > \mathbb{E}(t) \]
  - if ratio is decreasing, the stereotype is $\{t_1, \ldots, t_d\}$, and
    \[ \mathbb{E}^{st}(t|G) < \mathbb{E}(t|G) < \mathbb{E}(t) \]

- Mean assessment is shifted in the correct direction, but...

- Neglecting non-stereotypical tail leads to “over-confidence”:
  - DM’s estimate of the mean of $t$ is too extreme and different
  - DM’s assessment of the variance of $t$ is dampened
Extreme Stereotypes

Example: Income distribution in the US

- Stereotypes truncate left tail for Whites, right tail for Blacks
  - Survey respondents exaggerate the proportion of poor among Blacks (Gilens 1996)
  - Also the proportion of Blacks among poor (base rate neglect)
Extreme Stereotypes

This logic implies overreaction to diagnostic tests

- Example: investor assesses mutual fund manager’s skill level.

<table>
<thead>
<tr>
<th></th>
<th>high skill</th>
<th>no skill</th>
</tr>
</thead>
<tbody>
<tr>
<td>high returns</td>
<td>0.1</td>
<td>0.4</td>
</tr>
<tr>
<td>low returns</td>
<td>0.05</td>
<td>0.45</td>
</tr>
</tbody>
</table>

- $\Pr(\text{skill} | \text{high return}) / \Pr(\text{skill} | \text{low return})$: high returns generate high skill stereotype

- investor neglects that even managers with no skill can perform well due to luck

- Fund flows and past performance
Extreme Stereotypes (continuous case)

- Stereotype of $G$ is a truncation of $\pi_{t,G}$ to the most representative types whose aggregate probability mass is $\delta \leq 1$

$$T_G(c) = \left\{ t \in T \mid \frac{\pi_{t,G}}{\pi_{t,-G}} \geq c \right\}, \quad \int_{t \in T(c(\delta))} \pi_{t,G} dt = \delta$$

- Suppose $\pi_{t,W}$, are univariate normal, $\sigma_G = \sigma_{-G}$, $\mu_G > \mu_{-G}$
Extreme Stereotypes (continuous case)

- If \( \sigma_G < \sigma_{-G} \) (and \( \mu_G = \mu_{-G} \)) the stereotype of \( G \) is the central tendency
- If \( \sigma_G > \sigma_{-G} \) (and \( \mu_G = \mu_{-G} \)) the stereotype of \( G \) is the tails
- Examples:
  - Stocks are perceived as much more risky than corporate bonds
  - Immigrants are perceived either as high skilled, or as low skilled
Representativeness depends on $-G$.

- In the US, Whites are well off when compared to Blacks, but working class when compared to Asians
- Irish are red-haired when compared to Europeans, but Catholic when compared to Scotts

With Katie Coffman, we are developing experimental tests for the role of representativeness

- subjects briefly shown abstract groups $G$ and $-G$
- as $-G$ changes, compare estimates of:
  - modal type of $G$ (for unordered groups)
  - average type in $G$, $\mathbb{E}[t|G]$ (for ordered groups)
Preliminary Experimental Evidence

- Subjects observe two sets of 25 t-shirts for 15 secs
- Two conditions:
  - Control:
    - t-shirts | Girls (G) | Boys (−G_{Ctrl})
    - blue     | 0          | 0
    - green    | 12         | 12
    - purple   | 13         | 13
  - Representativeness:
    - t-shirts | Girls (G) | Boys (−G_{Rep})
    - blue     | 0          | 12
    - green    | 12         | 0
    - purple   | 13         | 13
- Then answer incentivized questions about observed sets
  - For each group (girls vs boys), what is modal color?
Control Condition
Representativeness Condition
Preliminary Experimental Evidence

- 202 subjects at Ohio State University Experimental Lab

![Bar chart showing the comparison between control treatment and representativeness treatment in correctly identified modal color and guessed that the "12 Shirt" color was modal.](chart.png)
Reaction to Information

How do stereotypes change as DM receives new information?

▶ Assume DM has priors over $\pi_{t,G}$, $\pi_{t,-G}$ such that:

$$\mathbb{E}(\pi_{t,W} | \alpha_W) = \frac{\alpha_{t,W}}{\sum_u \alpha_{u,W}}, \text{ for } W = G, -G$$

▶ Intuitively, prior counts the share of $t$ observations
  (Dirichlet priors are conjugate to Multinomial distribution)

▶ if new sample $n_W = (n_{1,W}, \ldots, n_{N,W})$ is observed,

$$\mathbb{E}(\pi_{t,W} | \alpha_W, n_W) = \frac{\alpha_{t,W} + n_{t,W}}{\sum_u (\alpha_{u,W} + n_{u,W})}$$

▶ Stereotype is determined from (Bayesian) $\mathbb{E}(\pi_{t,W} | \alpha_W, n_W)$. 
Reaction to Information

DM under-reacts to non-stereotypical information, provided it does not change the stereotype

- Suppose observation $t'$ outside the support of stereotype $\pi_{t,G}^{st}$
  - e.g. a successful member of a disadvantaged group
  - a bad year for a successful money manager

- If representativeness ranking for $G$ does not change, then probability distribution $\pi_{t,G}^{st}$ does not change
  - Non-stereotypical data coded as an “exception to the rule”

- Non-stereotypical information can be ineffective at changing beliefs even if it swamps instances in the stereotype
  - Confirmation bias
Reaction to Information and Stereotype Change

The stereotype for $G$ changes only if the new observations are sufficiently contrary to the initial stereotype

- $t$ replaced by $t'$ in support of stereotype when \[ \frac{\pi_t,G}{\pi_t,-G} < \frac{\pi_{t'},G}{\pi_{t'},-G} \]

- observe $t$ from group $-G \Rightarrow$ decreases \[ \frac{\pi_t,G}{\pi_t,-G} \]

- observe $t'$ from group $G \Rightarrow$ increases \[ \frac{\pi_{t'},G}{\pi_{t'},-G} \]

- Intuition: only asymmetric changes in groups are effective at changing stereotypes

- raising income of all groups does not change stereotype of a group that has disproportionate share of poor

- rise of successful commercial class from disadvantaged group changes stereotype for that group

- a run of above-market performance drastically improves money manager’s stereotype (and expectations of future performance)
Modeling Financial Crises

- Conventional approach: low probability bad outcome that materializes (MIT shock)

- Problems with this approach:
  - Crises are rather frequent (Reinhart and Rogoff)
  - Evidence that investors underestimate risks (Coval et al, Foote et al, Greenwood-Shleifer)

- Show that stereotypes can be used to build a psychologically founded theory of boom-bust beliefs and financial crises
In line with previous (and KT) notion of representativeness, we posit: a cash flow is representative of a sample of data if after observing the data the cash flow is much more likely to occur

- Formalization as in GS (2010) and BGS (2014)
- Link to notion of similarity

Many results follow: excess optimism after good news, under-reaction to occasional bad news, over-reaction to pattern of bad news (or news about low probability disasters)
A Model of Debt Issuance

- Three periods $t = 0, 1, 2$
- One asset, on which debt is issued at $t = 1$
  
  at time $t$, beliefs of terminal cash flow = \[
  \begin{cases}
  y_h & \text{with prob. } \pi^t_h \\
  y_l & \text{with prob. } \pi^t_l
  \end{cases}
  \]

- Investors’ preferences: risk neutral for probability of default $< \rho$, infinitely risk averse otherwise

- Profit (and debt) maximization: $\max \mathbb{E} [y - d + p(d)]$
Path of News and Bayesian Beliefs

- As before, the prior distribution is Dirichlet

- Between $t = 0$ and $t = 1$ observe $(n^0_h, 0)$ pieces of good news. Bayesian beliefs at $t = 1$:

$$\pi^1_h = \frac{\pi^0_h + n^0_h}{1 + n^0_h} > \pi^0_h.$$ 

- Between $t = 1$ and $t = 2$ observe news $(n^1_h, n^1_l)$. Bayesian beliefs at $t = 2$ are:

$$\pi^2_h = \frac{\pi^0_h + n^0_h + n^1_h}{1 + n^0_h + n^1_h + n^1_l} > \pi^0_h.$$
Rational Expectations

- Assume that $\rho < \pi^1_l$

- Then:
  - Debt issued at $t = 1$ is $d_1 = y_l$
  - Price of debt at $t = 1$ is $p(d_1) = y_l$
  - Price does not change after bad news at $t = 2$
At $t$, the representativeness of cash flow $y_k$ is:

$$R(y_k, t) = \frac{\pi^t_k}{\pi^{t-1}_k}.$$  

The ”group” $G$ is: terminal cash flow as of time $t$, after news are received. Compare it to $-G$: terminal cash flow as of $t - 1$, the pre-news period.

Investors deflate by factor $\delta \in [0, 1]$ the less representative cash flow.

Idea: inflate the odds of cash flows that have become relatively more likely in light of the data.
Debt at $t=1$

- After observing $n_h^0$ good news, $y_h$ is representative:
  - The investor assesses
    \[
    \pi_i^r,1 = \frac{\pi_i^0 \delta}{(\pi_h^0 + n_h^0) + \pi_i^0 \delta} < \pi_i^1.
    \]
  - For $\delta$ small enough we have $\rho > \pi_i^r,1$, so that debt issuance/prices are:
    \[
    d_i^r = y_h, \quad p^r(d_i^r) = y_h - \pi_i^r,1(y_h - y_l).
    \]
- The Investor is too optimistic. Debt goes up relative to RE.
When $\delta \to 0$, there is no perceived downside outcome, the maximal cash flow is pledged, and the full price is extracted:

$$\pi_{r,1}^l = \frac{\pi_0^l \delta}{(\pi_h^0 + n_h^0) + \pi_l^0 \delta} \to 0$$

$$d^r_1 = y_h$$

$$p^r(d^r_1) \to y_h$$

Two remarks:

- "This time is different syndrome": crash is neglected even though it is quite likely
- Implication on reaction to the data: level versus change
Debt at t=2

- After observing \((n^1_h, n^1_i)\) mixed news, for \(\delta \to 0\) we have that:
  - If \(n^1_i/n^1_h < \pi^0_i/(\pi^0_h + n^0_h)\), the high cash flow remains representative and the investor under-reacts to information, namely \(\pi^{r,2}_i = 0\)
  - If \(n^1_i/n^1_h > \pi^0_i/(\pi^0_h + n^0_h)\), the low cash becomes representative and the investor over-reacts to information, namely \(\pi^{r,2}_i = 1\)

- A few pieces of bad news are ignored, the investor over reacts to systematically bad news
  - When representativeness changes, past info is re-evaluated
  - Over-reaction to news about unlikely events
To Conclude

- Representativeness-based model of stereotype formation and change
- The logic of stereotypes can account for ex-ante neglect of risk and ex-post exaggeration of it as a function of the data
- Representativeness generates excess volatility in beliefs and boom-bust cycles in the absence of standard amplifiers
- Unify neglect of risk, under-reaction and over-reaction in a psychologically founded model that rationalized many puzzles from lab and field
  - confirmation bias
  - base rate neglect
- Model is highly tractable and can be applied to many other domains (e.g. identity)
To Conclude

Open questions:

- Recall of types may also be driven by availability or likelihood (as in categorisation models, Mullainathan 2002, Freyer and Jackson 2007)
- Do stereotypes feature an in-group vs out-group asymmetry?
- In problems with less structure, what determines \(-G, T\)?