Modeling Movie Choices and Ratings

Tim Rubin,
Mark Steyvers

Outline

- Recommendation systems
  - Overview
  - Relevance to cognitive science
- Present a model for movie choices and ratings
  - Grounded in psychology
  - Interpretable dimensions of preference
- Using user choices to predict ratings

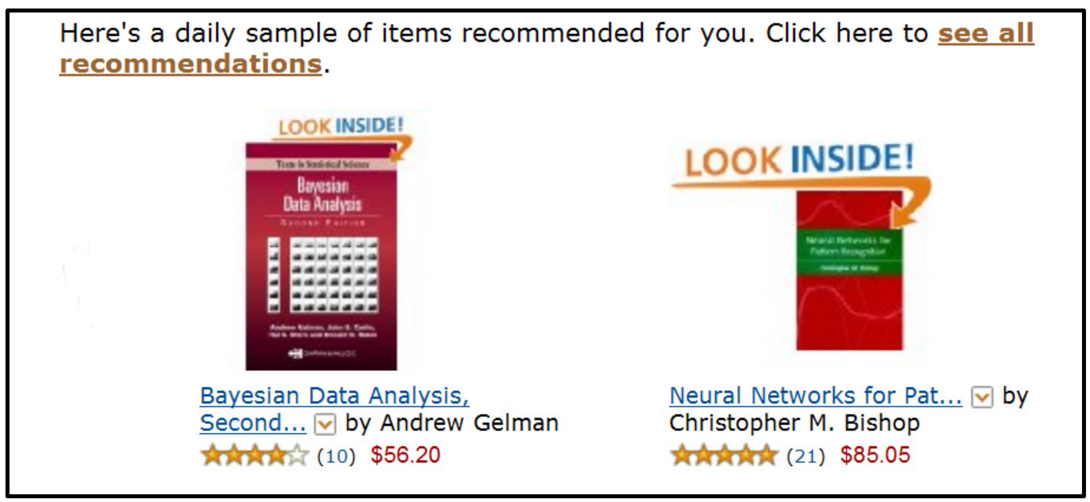
Recommendation Systems

- What is a recommendation system?
  - Based on user data, suggest new items to users
    - What will they *buy*?
    - **What will they *like*?**
- This research focuses on the second question

Examples of Recommendation Systems

- Netflix

- Amazon



Recommendation and Cognitive Science

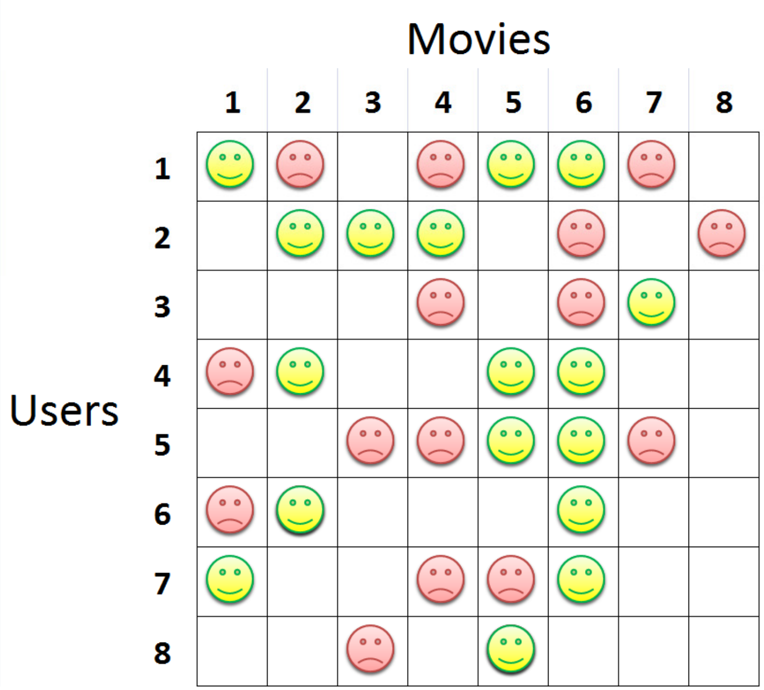
- What is the process that generates user input?
- E.g., a Netflix rating:
  - Choose a movie
  - Form an opinion
  - Choose a rating to reflect opinion

Recommendation and Cognitive Science

- Types of Input
  - Explicit: Ratings
  - Implicit: Choices

Two Views of Netflix Data

MOVIE RATINGS



MOVIE CHOICES

Movies

Users	1	2	3	4	5	6	7	8
1	✓	✓		✓	✓	✓	✓	
2		✓	✓	✓		✓		✓
3				✓		✓	✓	
4	✓	✓			✓	✓		
5			✓	✓	✓	✓	✓	
6	✓	✓				✓		
7	✓			✓	✓	✓		
8			✓		✓			

Types of input

- Suppose you know the following:

Choices	**Ratings**
The Godfather	?
Scarface	?
Full Metal Jacket	?

Types of input

- Suppose you know the following:

Choices	**Ratings**
The Godfather	☆☆☆☆
Scarface	☆☆☆
Full Metal Jacket	☆☆☆☆☆

- How much more do we know about this users preferences?

Current state of recommender systems

- Based on predicting ratings for missing items
- Popular methods
  - K-Nearest Neighbors (kNN)
    - Find similar users / movies
  - Singular Value Decomposition (SVD)
    - Spatial representation for users / movies

Problems with standard approaches

- No Psychological Basis
  - Lack true models / generative processes
  - Don’t account for choice
- Problems for Recommendation
  - Only use explicit data (ratings)
  - Prediction vs. recommendation

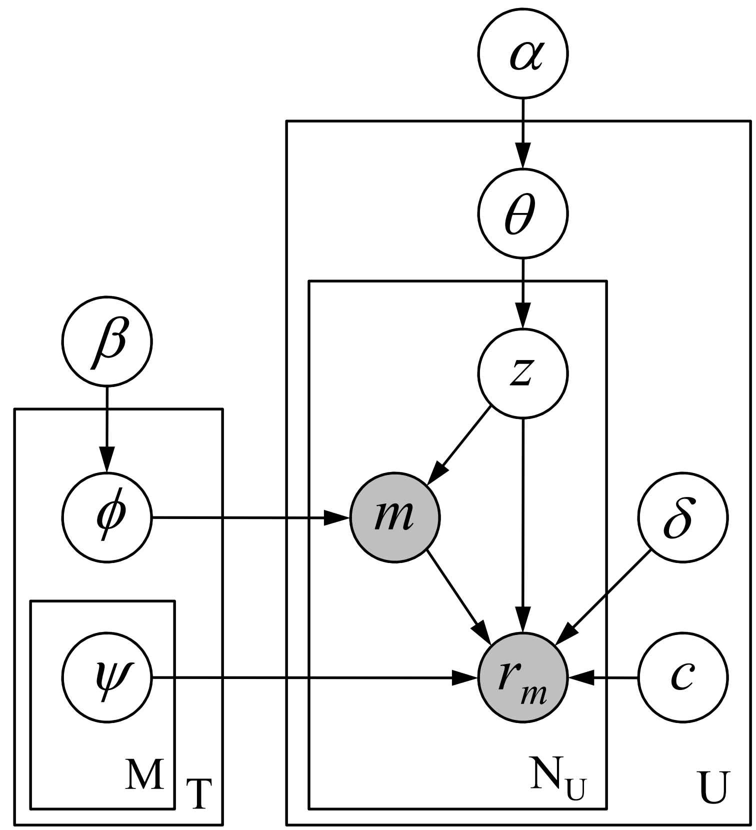
Research Goals

- Building a model based in psychology
  - A generative model for both movie-choices and movie ratings
- Explicit vs. Implicit Data
  - Can implicit data (*choices*) be used to improve predictions about explicit data (*ratings*)?

The Ratings Topic Model

- Describes process by which users:
  1. Choose movies
  2. Reach an opinion of movies
  3. Select a rating to reflect opinion

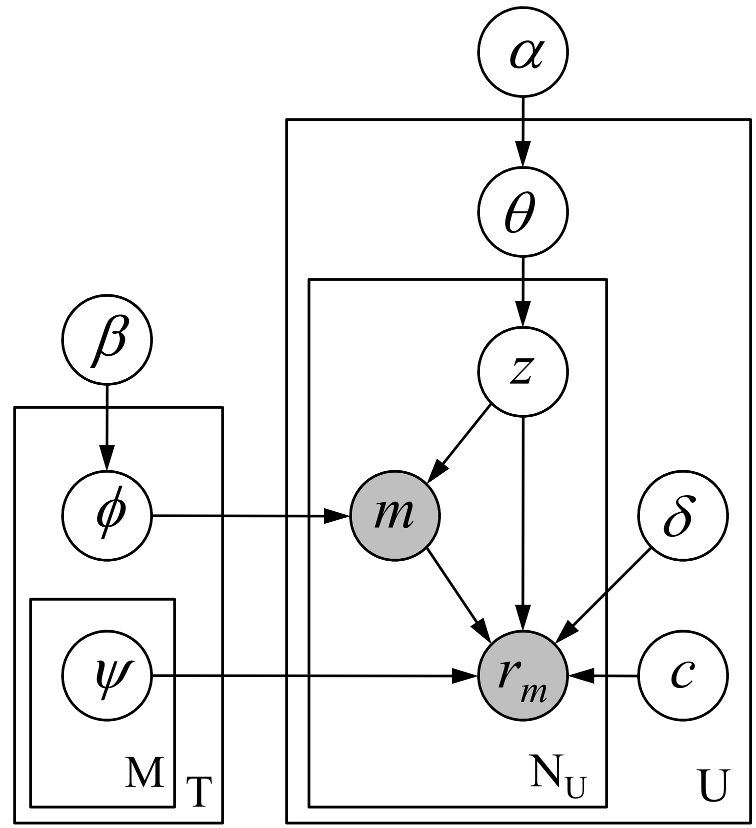
- Combines two established statistical methods
  1. Topic modeling
  2. Ordered-logit model



The Ratings Topic Model

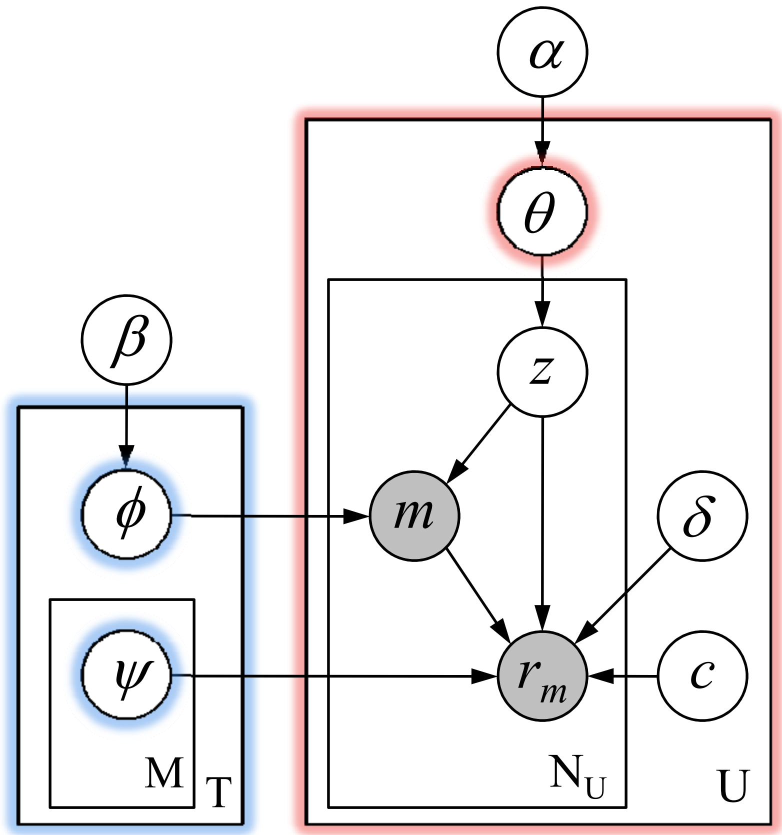
- Model trained using dense subset of Netflix dataset
- 10,000 users x 500 movies*
- Parameters learned using MCMC methods

***Note: consists of both movies *and* TV shows**



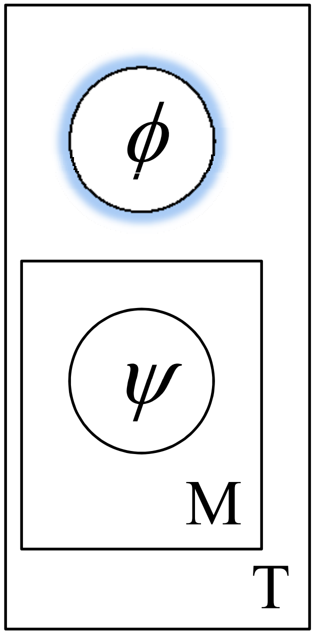
The Ratings Topic Model

- Users
- Topics
- Mixture Model
  - Users are mixtures of topics (θ)
- Each topic describes:
  - Movie choices (ϕ)
  - Movie preferences (ψ)



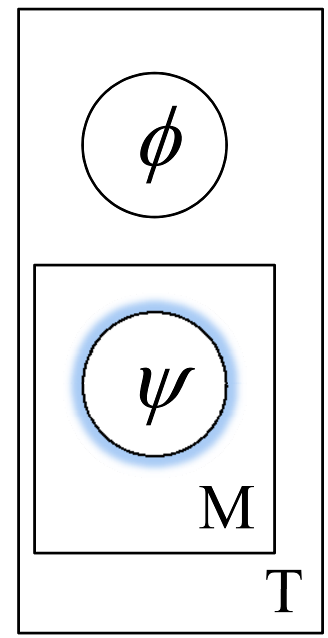
Topic Examples: Choice (ϕ)

Topic 4		Topic 22	
p **(*m*|*t*)**	**Most Likely Choices**	***p*** **(*m*|*t*)**	**Most Likely Choices**
.031	Poltergeist	.030	North by Northwest
.030	Carrie	.028	The Great Escape
.029	A Nightmare on Elm Street	.028	The Maltese Falcon
.027	Halloween	.028	Vertigo
.025	Misery	.028	The Bridge on the River Kwai
.024	Scream	.027	Some Like It Hot
.023	Saw	.025	Mr. Smith Goes to Washington
.022	The Exorcist	.025	Lawrence of Arabia
.022	The Grudge	.024	Cool Hand Luke
.021	The Lost Boys	.024	12 Angry Men
.021	Friday the 13th	.023	Butch Cassidy and the Sundance Kid
.020	Final Destination 2	.023	The Manchurian Candidate
.020	Stir of Echoes	.023	A Streetcar Named Desire
.020	Sleepy Hollow	.022	On the Waterfront
.019	Frailty	.021	All About Eve
.017	From Hell	.020	Patton
.017	I Know What You Did Last Summer	.018	All the President's Men
.016	The Haunting	.017	West Side Story
.016	Rosemary's Baby	.017	Bonnie and Clyde

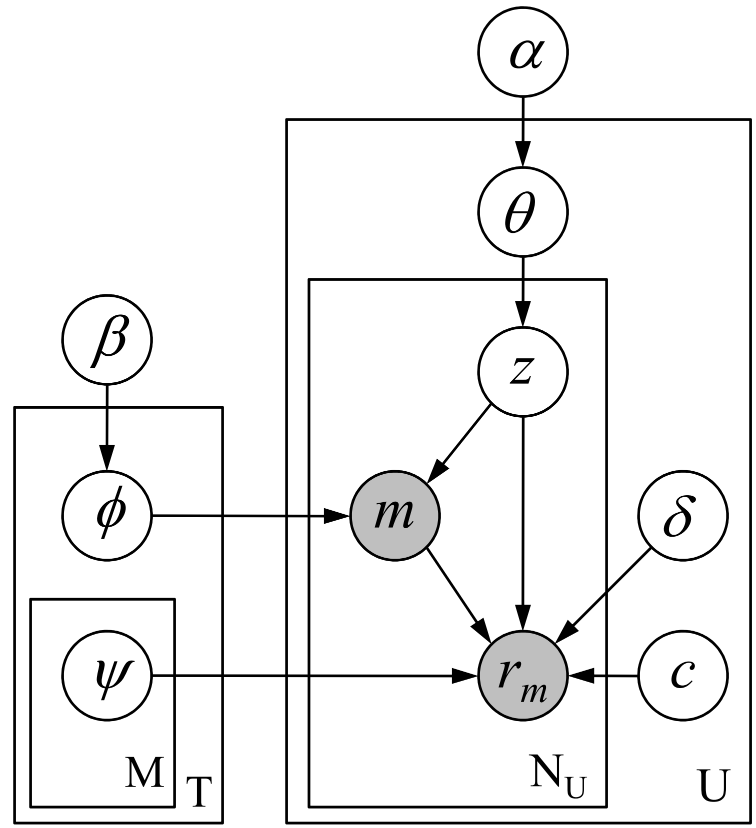


Topic Examples: Preference (ψ)

Topic 16		Topic 18	
ψ	**Most Enjoyed Movies**	ψ	**Most Enjoyed Movies**
3.31	The Sting	5.54	Band of Brothers
3.06	Fiddler on the Roof	5.52	The Sopranos: Season 1
3.05	12 Angry Men	5.35	24: Season 1
3.05	West Side Story	5.28	The Sopranos: Season 2
2.84	Gandhi	5.19	The Sopranos: Season 3
2.83	Butch Cassidy and the Sundance Kid	5.17	The Sopranos: Season 4
2.79	My Fair Lady: Special Edition	3.94	Hoosiers
2.77	Band of Brothers	3.89	Glory
2.71	The Great Escape	3.74	Casino: 10th Anniversary Edition
2.71	Moonstruck	3.72	Swingers
2.65	The King and I	3.51	The Sting
2.64	Mr. Smith Goes to Washington	3.50	The Natural
ψ	**Most Disliked Movies**	ψ	**Most Disliked Movies**
-1.84	Waiting for Guffman	-2.19	Shaun of the Dead
-1.87	Monster-in-Law	-2.28	Bring It On
-1.96	Before Sunset	-2.58	Batman & Robin
-2.06	Y Tu Mama Tambien	-2.65	Envy
-2.67	Halloween	-2.82	Little Nicky
-2.77	A Nightmare on Elm Street	-2.91	Alexander: Director"s Cut
-3.05	I Heart Huckabees	-3.63	White Chicks

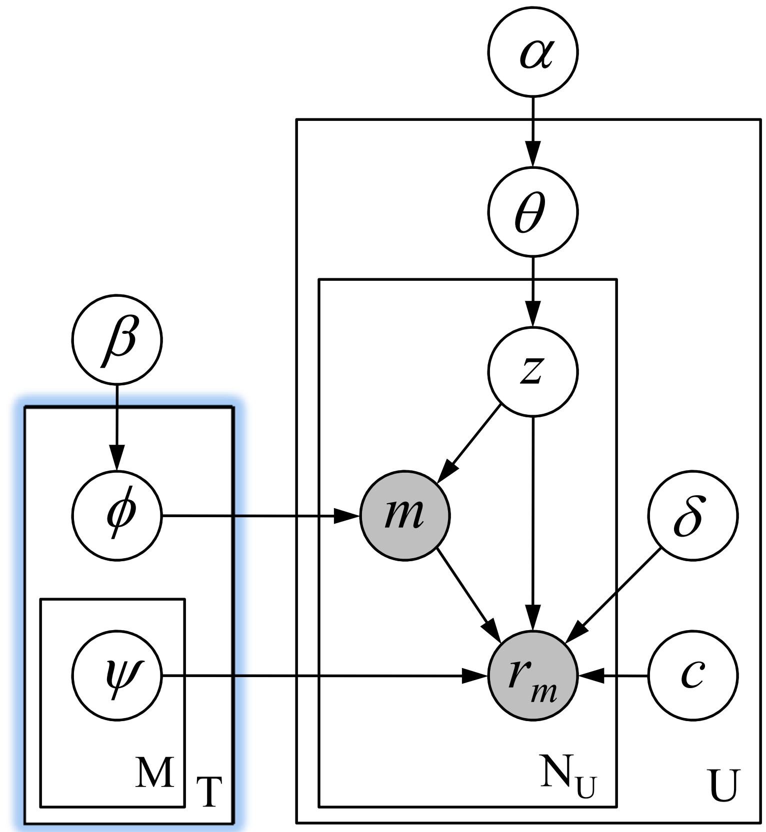


Generative Model



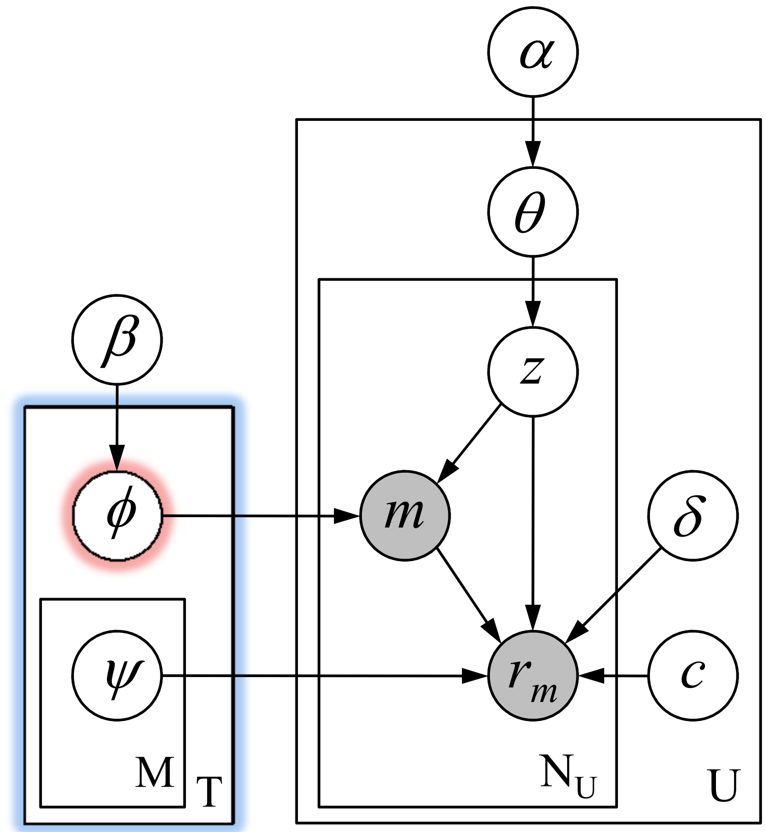
Generative Model

- **For each topic:**



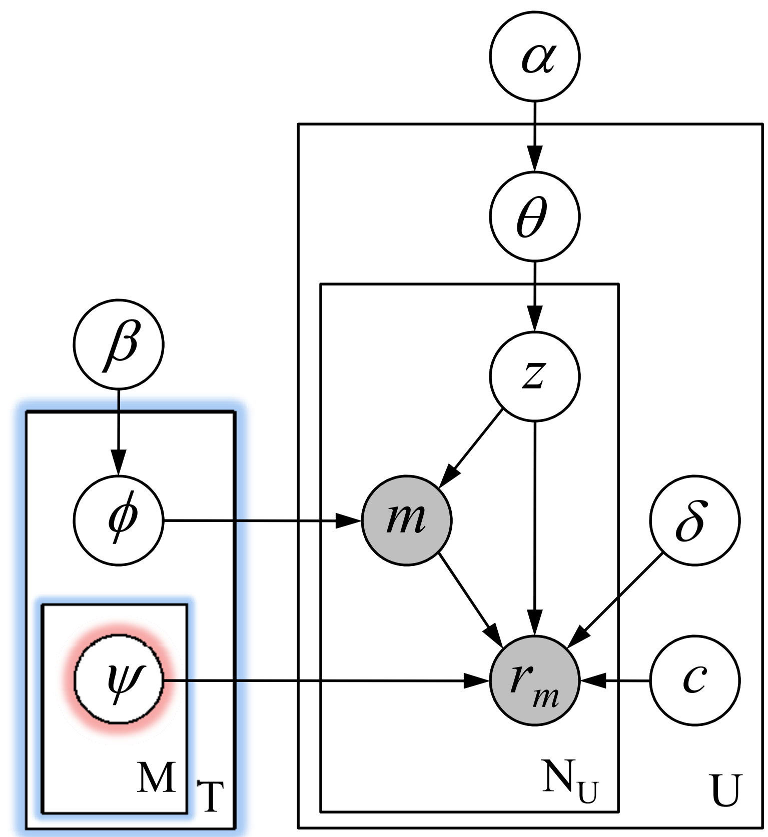
Generative Model

- **For each topic:**
  - Assign a multinomial distribution over movies (ϕ)



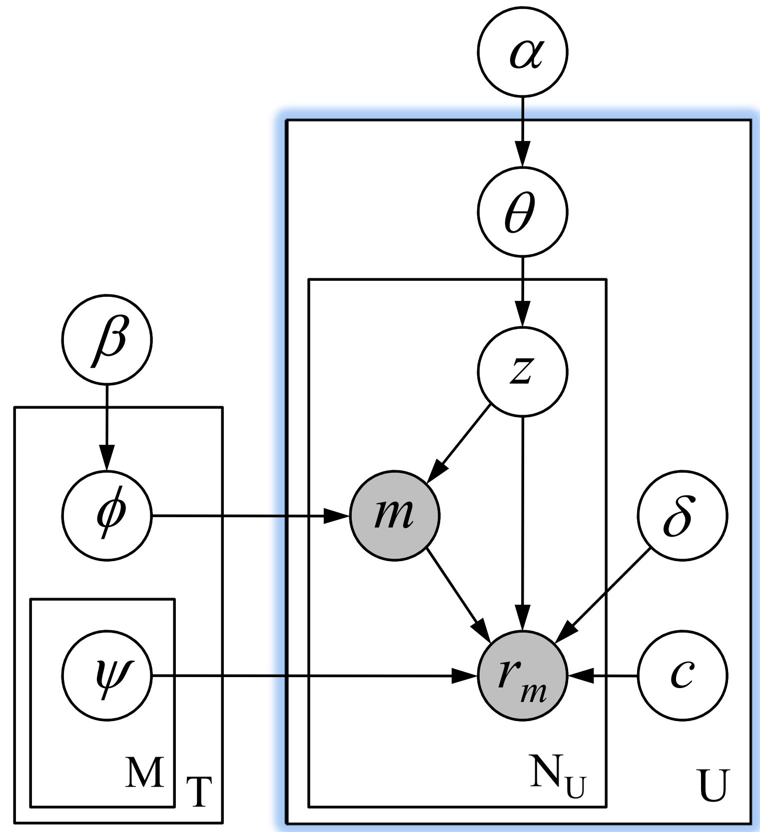
Generative Model

- **For each topic:**
  - Assign a multinomial distribution over movies (ϕ)
  - For each movie, assign a preference parameter (ψ)



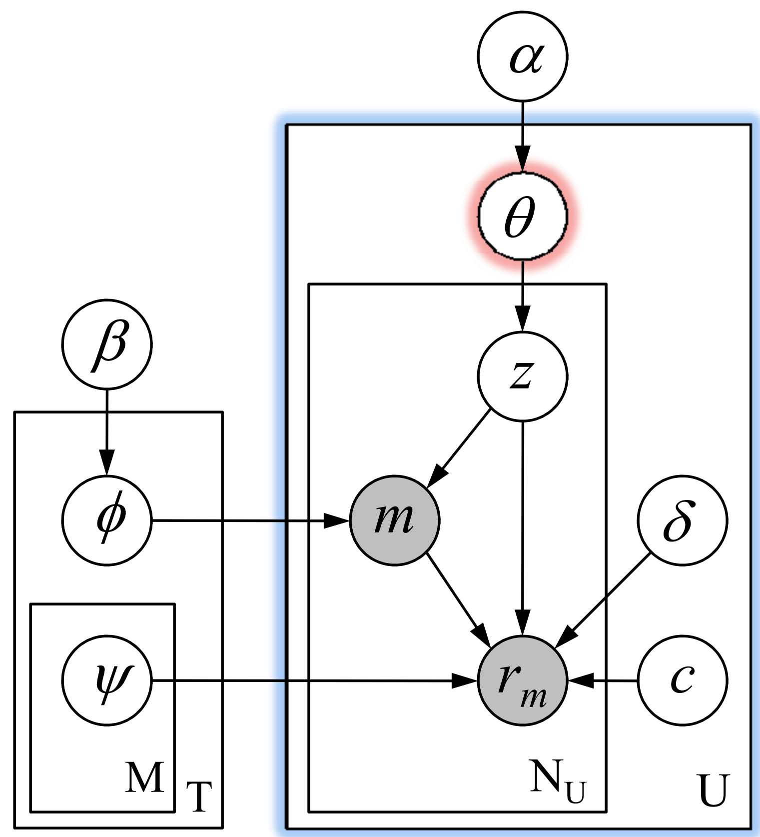
Generative Model

- For each topic:
  - Assign a multinomial distribution over movies (ϕ)
  - For each movie, assign a preference parameter (ψ)
- **For Each User:**



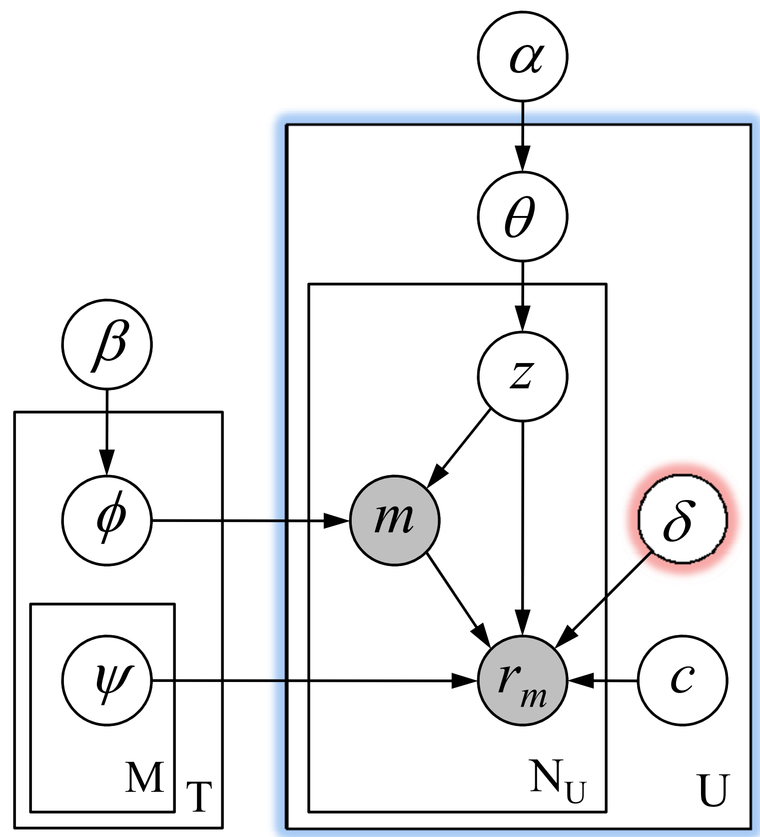
Generative Model

- For each topic:
  - Assign a multinomial distribution over movies (ϕ)
  - For each movie, assign a preference parameter (ψ)
- **For Each User:**
  - Assign a multinomial distribution over topics (θ)



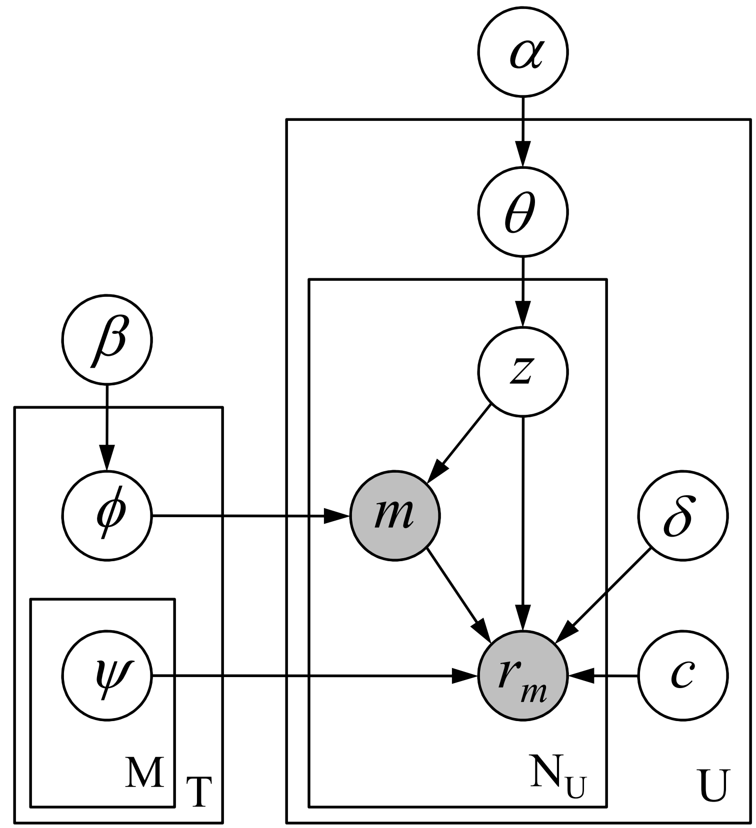
Generative Model

- For each topic:
  - Assign a multinomial distribution over movies (ϕ)
  - For each movie, assign a preference parameter (ψ)
- **For Each User:**
  - Assign a multinomial distribution over topics (θ)
  - Assign a rating-bias parameter (δ)



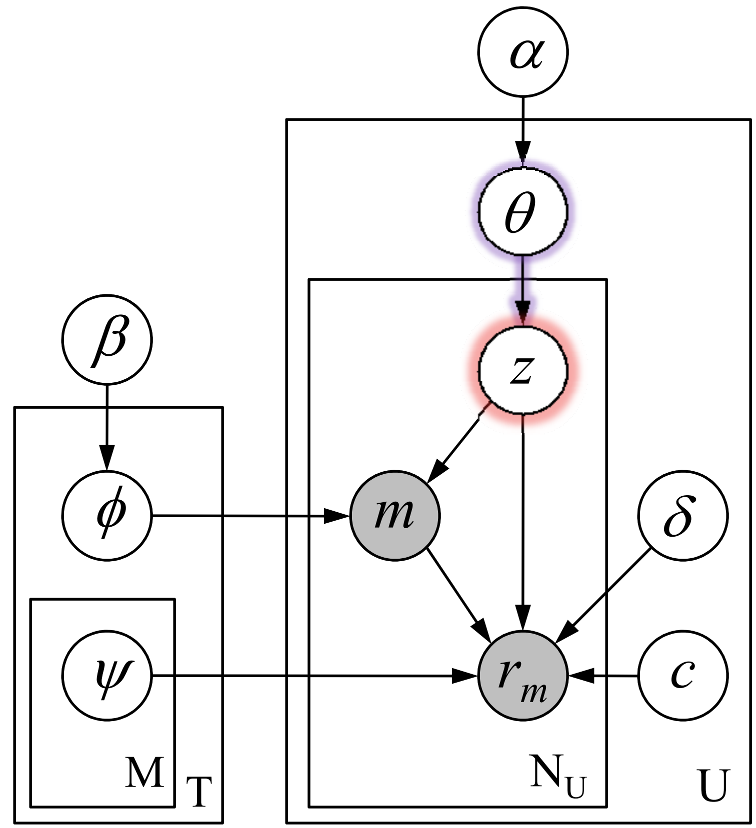
Generative Model

- For each topic:
  - Assign a multinomial distribution over movies (ϕ)
  - For each movie, assign a preference parameter (ψ)
- For Each User:
  - Assign a multinomial distribution over topics (θ)
  - Assign a rating-bias parameter (δ)
- **To generate a rating:**



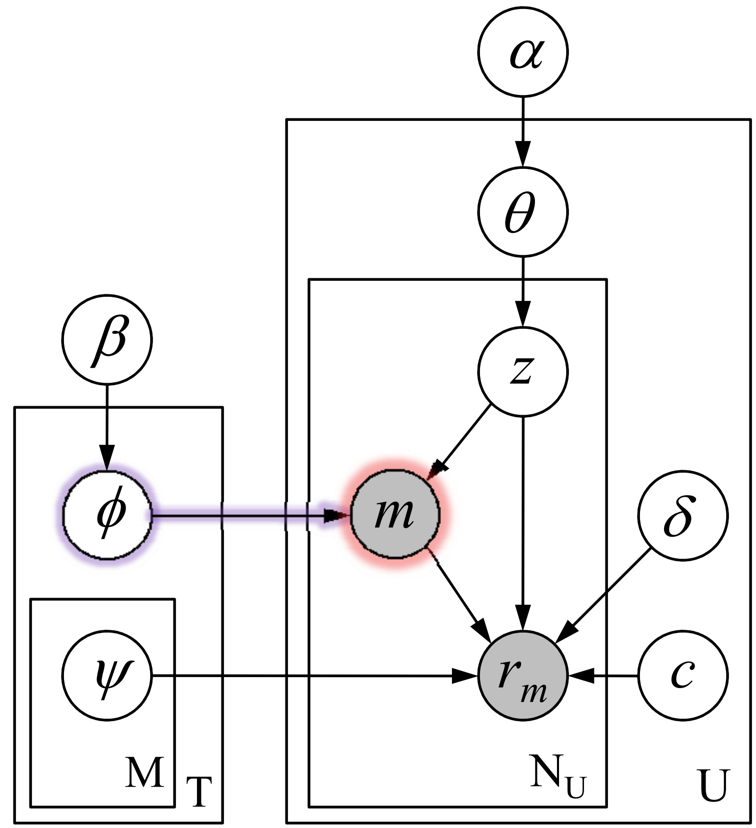
Generative Model

- For each topic:
  - Assign a multinomial distribution over movies (ϕ)
  - For each movie, assign a preference parameter (ψ)
- For Each User:
  - Assign a multinomial distribution over topics (θ)
  - Assign a rating-bias parameter (δ)
- **To generate a rating:**
  - Sample a topic (*z*) from a user's mixture



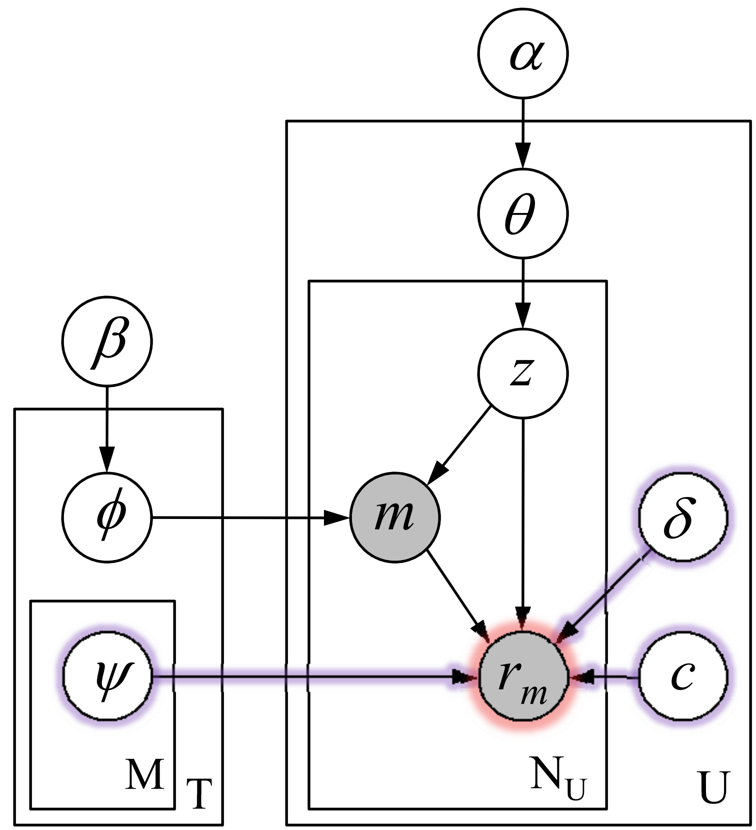
Generative Model

- For each topic:
  - Assign a multinomial distribution over movies (ϕ)
  - For each movie, assign a preference parameter (ψ)
- For Each User:
  - Assign a multinomial distribution over topics (θ)
  - Assign a rating-bias parameter (δ)
- **To generate a rating:**
  - Sample a topic (*z*) from a user's mixture
  - Sample a movie from topic *z*



Generative Model

- For each topic:
  - Assign a multinomial distribution over movies (ϕ)
  - For each movie, assign a preference parameter (ψ)
- For Each User:
  - Assign a multinomial distribution over topics (θ)
  - Assign a rating-bias parameter (δ)
- **To generate a rating:**
  - Sample a topic (*z*) from a user's mixture
  - Sample a movie from topic *z*
  - Use the ordered-logit model to translate preference into a rating



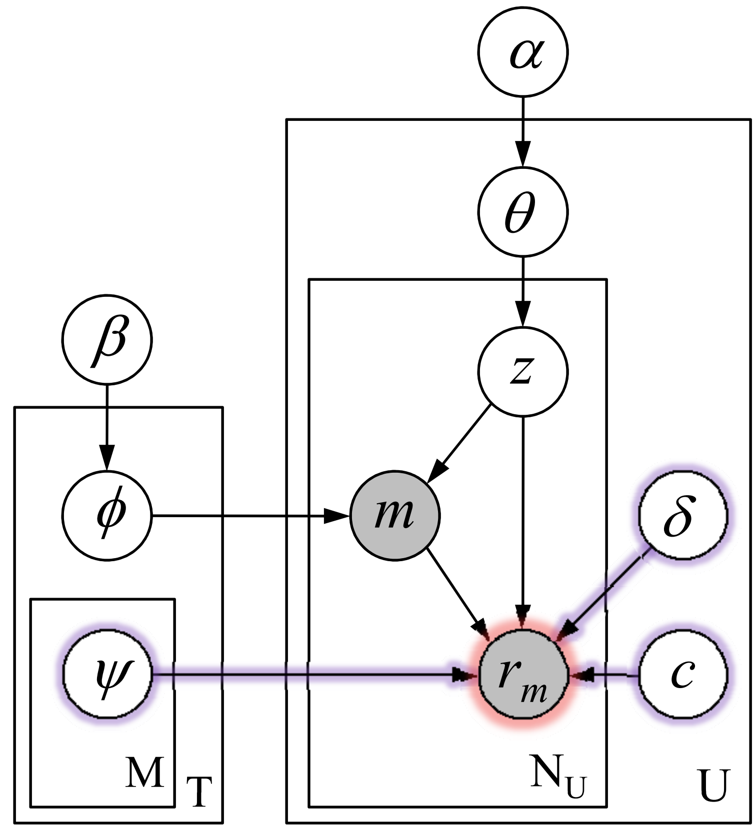
Ordered-Logit Model

- Converts preference into a discrete rating

ψ : Movie preference

δ : Rating bias

c : Ratings thresholds

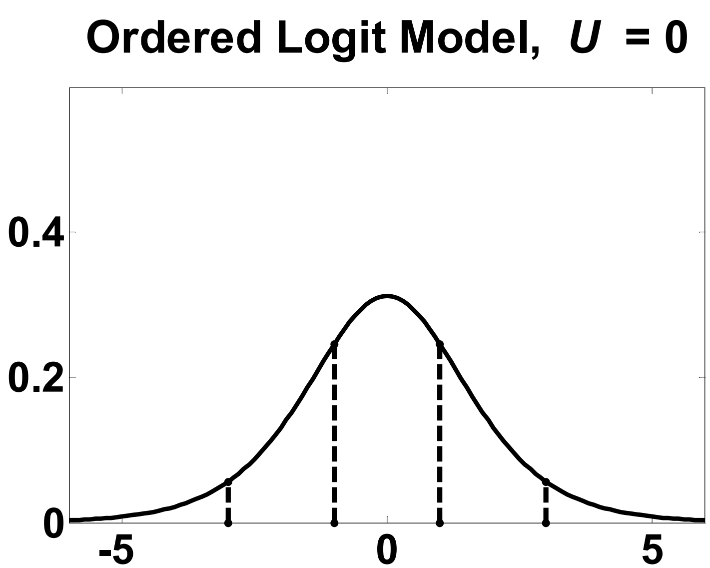


Ordered-Logit Model

- Rating probabilities are a function of utility

$$U_{u,m} = \psi_{t,m} + \delta_u + \varepsilon$$

- Cutoffs (c) define rating thresholds



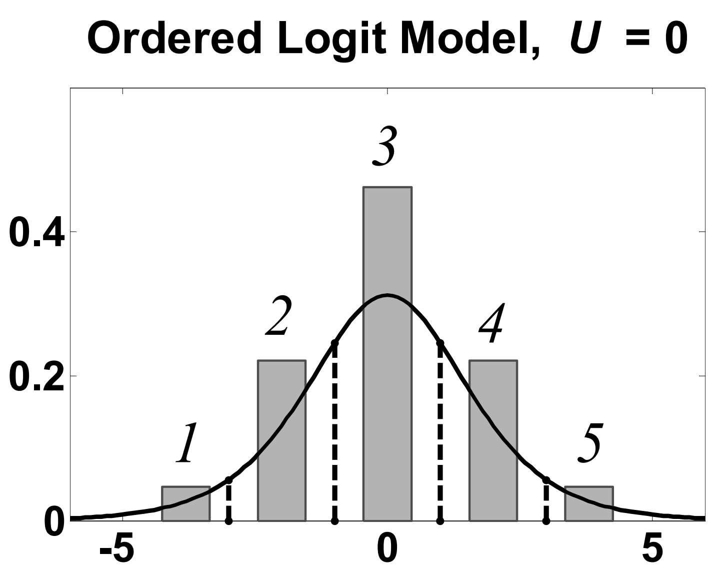
Ordered-Logit Model

- Rating probabilities are a function of utility

$$U_{u,m} = \psi_{t,m} + \delta_u + \varepsilon$$

- Cutoffs (c) define rating thresholds

$$P\left(r = r_{u,m} \mid \psi_m, \delta_u\right) = P(c_i < U_{u,m} < c_{i+1})$$



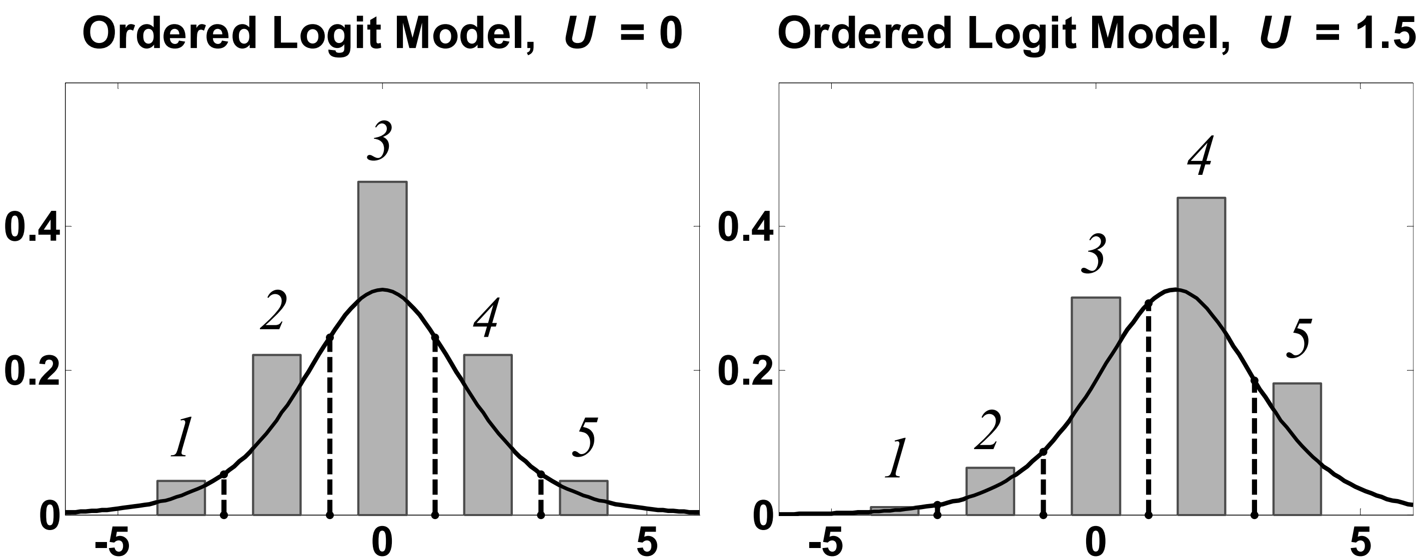
Ordered-Logit Model

- Rating probabilities are a function of utility

$$U_{u,m} = \psi_{t,m} + \delta_u + \varepsilon$$

- Cutoffs (c) define rating thresholds

$$P\left(r = r_{u,m} \mid \psi_m, \delta_u\right) = P(c_i < U_{u,m} < c_{i+1})$$



Predicting User Ratings

- Probability that user *u* gives rating *v* to movie *m:*

$$p(r=v\mid u,m) \propto \sum_{z=1}^{T} \underbrace{p(r=v\mid m,u,z)}_{\text{Ordered-Logit Model}}\ \underbrace{p(m\mid z)\ p(z\mid u)}_{\text{Mixture Model}}$$

An Example Topic

- What can we learn from each topic?
- Relevance to recommendation

Choice Dimension $p(m \mid t)$		Preference Dimension $E(r \mid m, t)$		Joint Probability $p(r, m \mid t)$
		Topic 4		
p	***Most Likely Choices***	**E(*r*)**	***Highest Rated***	***Most Likely To Please***
.031	Poltergeist	4.4	Labyrinth	The Exorcist
.030	Carrie	4.2	The Exorcist	Poltergeist
.029	A Nightmare on Elm Street	4.2	The NeverEnding Story	Misery
.027	Halloween	4.2	Aliens	Halloween
.025	Misery	4.1	Alien	A Nightmare on Elm Street
.024	Scream	4.0	Primal Fear	Carrie
.023	Saw	4.0	Superman: The Movie	The Lost Boys
.022	The Exorcist	4.0	Misery	Scream
.022	The Grudge	4.0	Poltergeist	Saw
.021	The Lost Boys	4.0	South Park: Bigger, Long…	Alien
.021	Friday the 13th	4.0	Lean on Me	Bram Stoker's Dracula
.020	Final Destination 2	4.0	The Life of David Gale	Aliens
.020	Stir of Echoes	3.9	Bram Stoker's Dracula	Stir of Echoes
.020	Sleepy Hollow	3.9	Thelma & Louise	Frailty
.019	Frailty	3.9	Halloween	Dawn of the Dead
.017	From Hell	3.9	The Lost Boys	Labyrinth
.017	I Know What You Did La…	3.9	Sleepers	Fatal Attraction
.016	The Haunting	3.9	Hostage	The NeverEnding Story
.016	Rosemary's Baby			
.016	Hide and Seek	**E(*r*)**	***Lowest Rated***	***Most Likely To Dissapoint***
.016	Bram Stoker's Dracula	2.3	Where the Heart Is	Dreamcatcher
.016	Dreamcatcher	2.3	Dr. Dolittle 2	The Ring Two
.015	Stigmata	2.2	Sneakers	White Noise
.015	Resident Evil	2.2	Team America: World Pol…	The Haunting
.014	The Ring Two	2.1	The English Patient	Catwoman
.014	The Gift	2.1	Black Sheep	The Grudge
.014	Fatal Attraction	2.0	Catwoman	Hide and Seek
.013	Alien	1.9	8 Mile	Scary Movie 2

Prediction Tasks

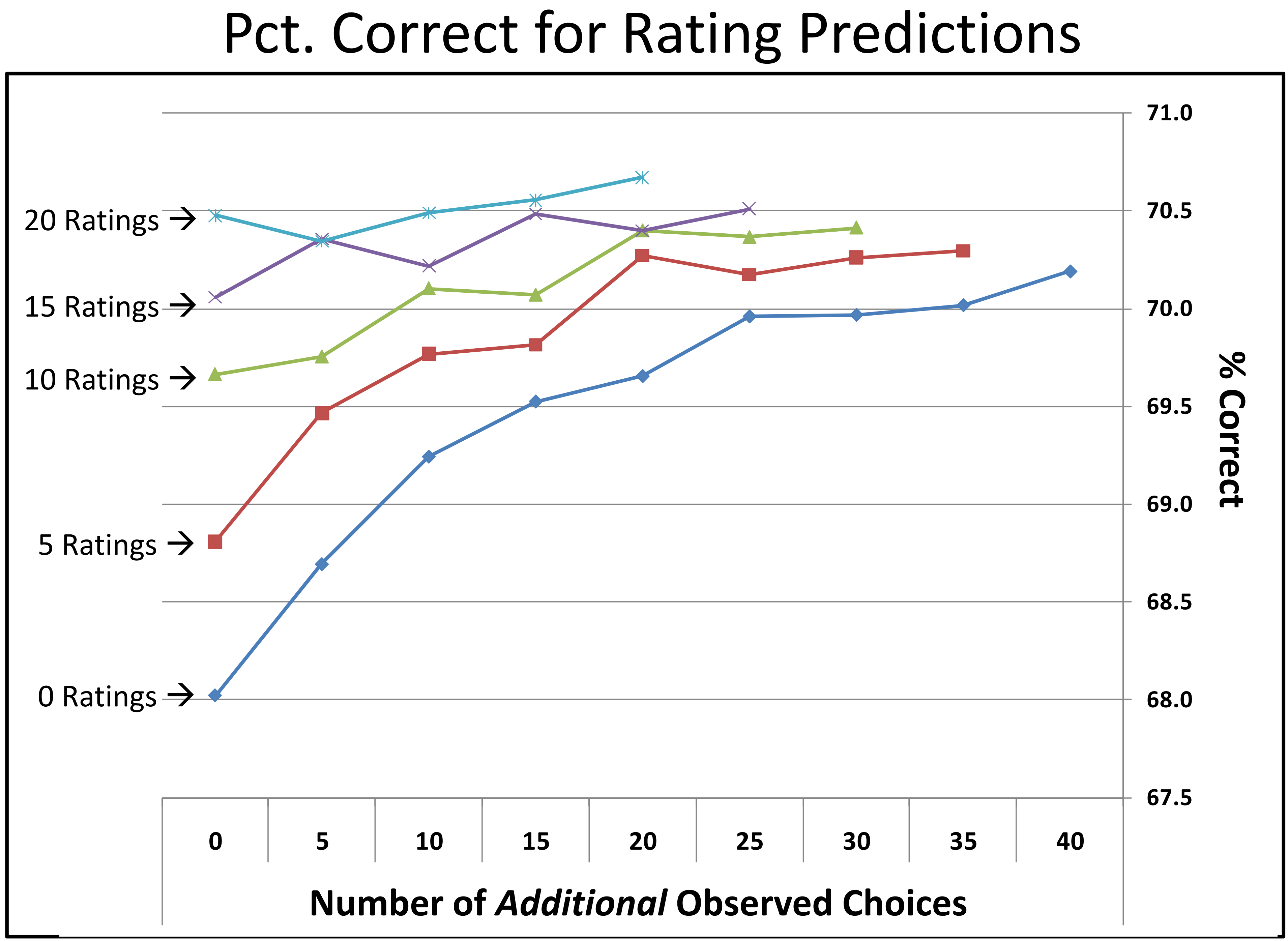
- Able to make predictions for withheld data:
  - User *ratings*
  - User *choices*
- For time purposes, will focus on question:
  - Can we use user *choice* to improve *rating* prediction?

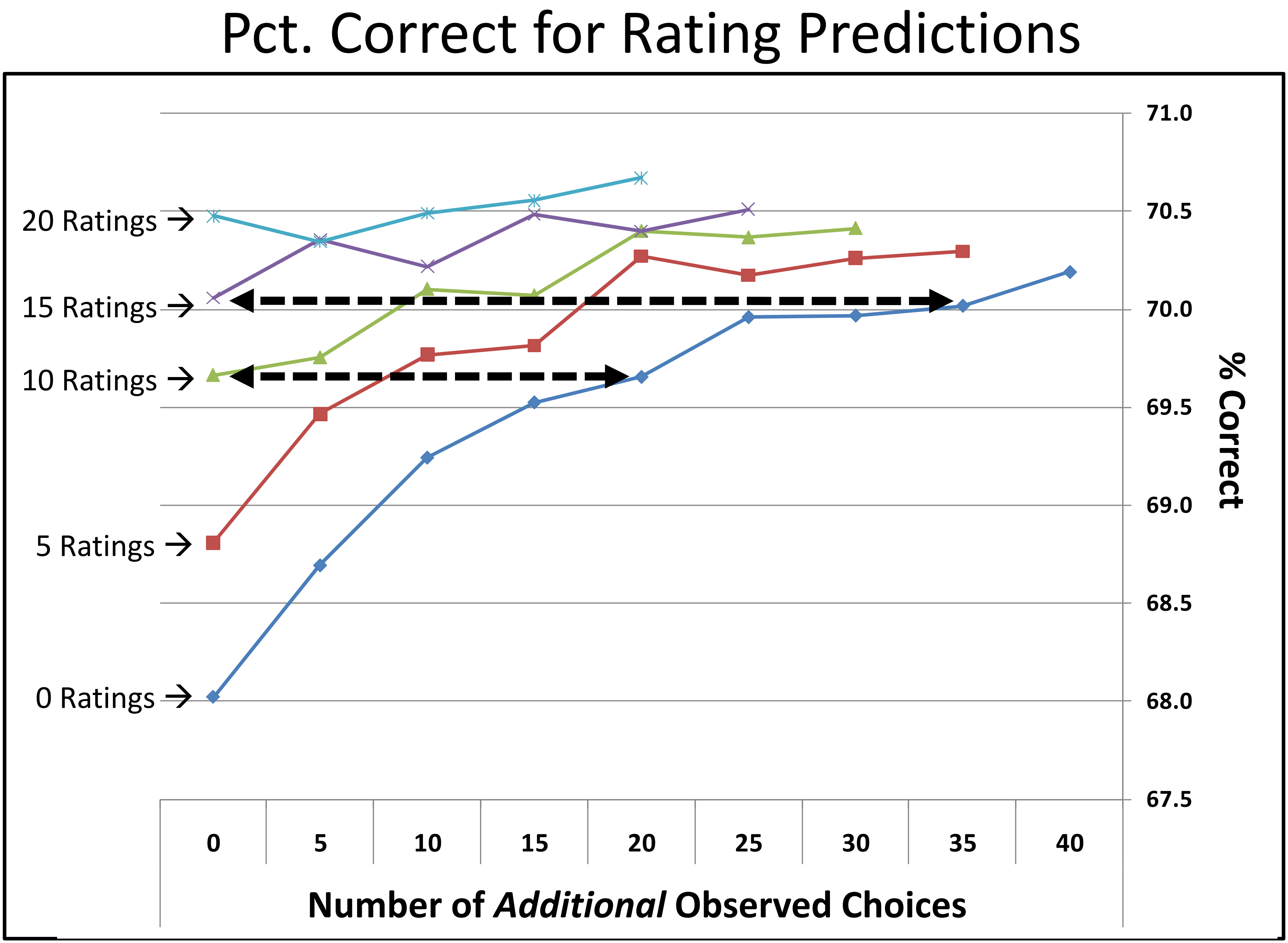
What can we learn from user choices?

- Trained topic-parameters on 10,000 users
- Evaluated model on 1,000 new "test users"
- Systematically manipulated number of observed:
  - Ratings (*explicit* data)
  - Choices (*implicit* data)

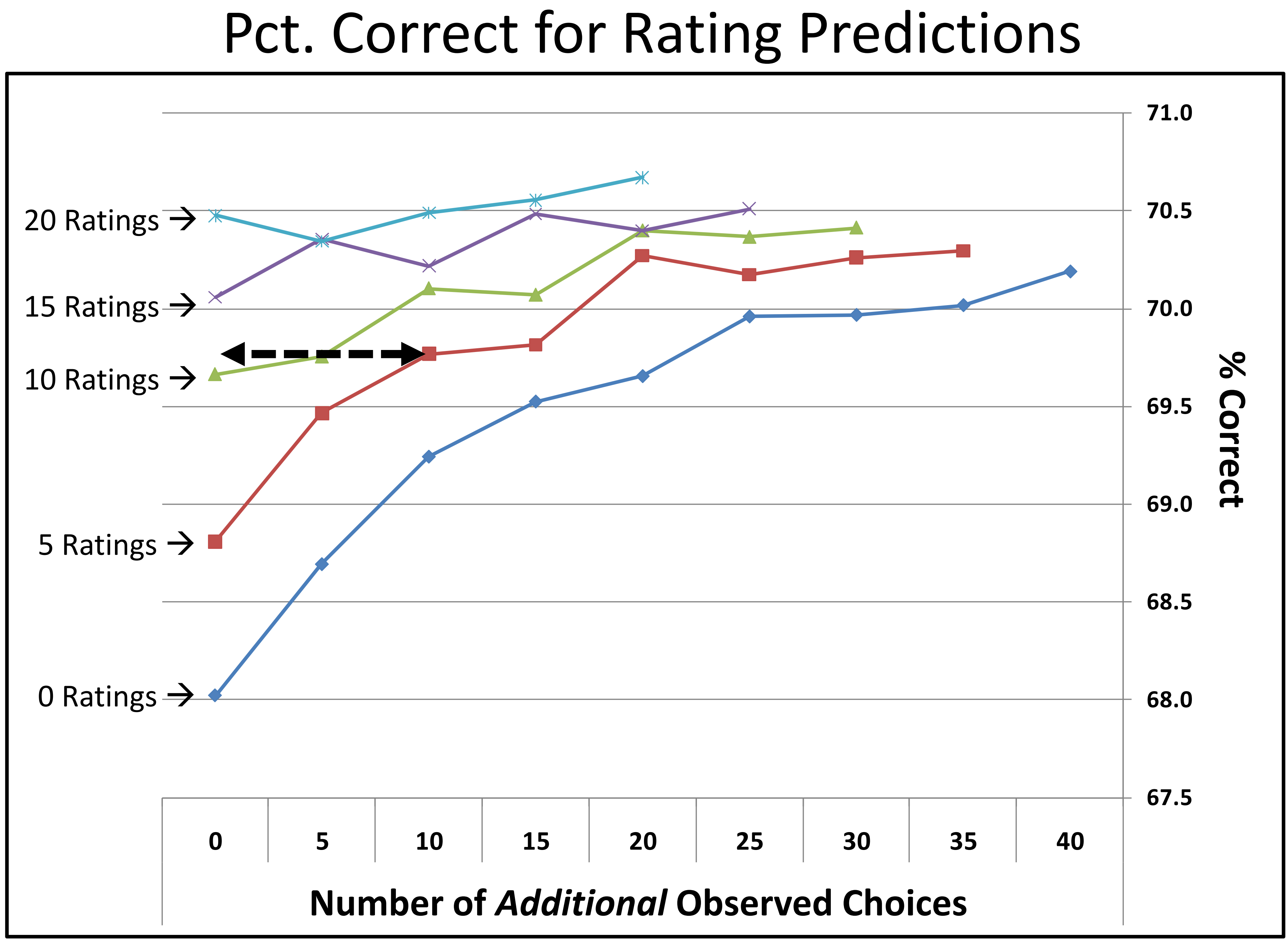
Performance Measure

- For all unequal pairs of withheld test ratings, predict which rating will be higher


Pct. Correct for Rating Predictions
20 Ratings
15 Ratings
10 Ratings
5 Ratings
0 Ratings
71.0
70.5
70.0
69.5
69.0
68.5
68.0
67.5
% Correct
0
5
10
15
20
25
30
35
40
Number of Additional Observed Choices



Pct. Correct for Rating Predictions
20 Ratings
15 Ratings
10 Ratings
5 Ratings
0 Ratings
71.0
70.5
70.0
69.5
69.0
68.5
68.0
67.5
% Correct
0
5
10
15
20
25
30
35
40
Number of *Additional* Observed Choices



Pct. Correct for Rating Predictions
20 Ratings
15 Ratings
10 Ratings
5 Ratings
0 Ratings
71.0
70.5
70.0
69.5
69.0
68.5
68.0
67.5
% Correct
0
5
10
15
20
25
30
35
40
Number of *Additional* Observed Choices

Summary

- Presented the Ratings Topic Model
  - Generative model for human choice processes and preferences
  - Captures interpretable dimensions of choice and preference
- Demonstrated that user choices can be used to improve predictions about preferences

Special Thanks

- Mark Steyvers
- Madlab