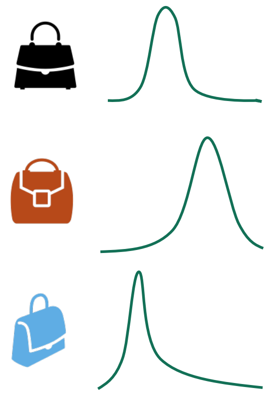
Forecasting Demand Distributions for New Products: Combining Sales Data with Subjective Rankings

Yale SCHOOL OF MANAGEMENT

Marat Salikhov

December 21st, 2020

$$Q^* = F^{-1}\left(\frac{c_u}{c_u + c_o}\right)$$




Quantity
190
130
230
180
120
100
2018
2019
Sales Data
+
2 1 3
2 3 1
1 3 2
Subjective Rankings
=
$Q^* = F^{-1}\left(\frac{c_u}{c_u + c_o}\right)$

Wisdom of crowds

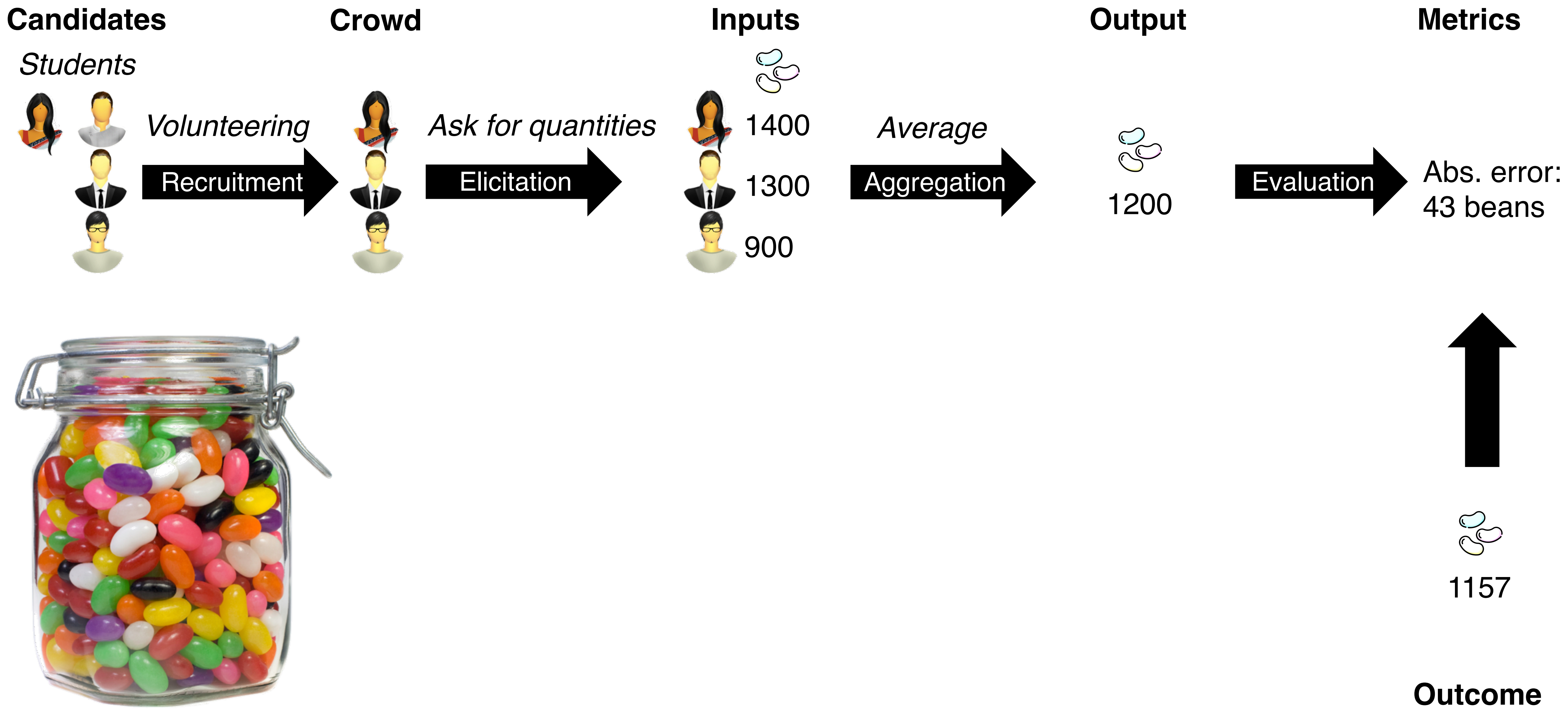
Candidates $\xrightarrow{\text{Recruitment}}$ **Crowd** $\xrightarrow{\text{Elicitation}}$ **Inputs** $\xrightarrow{\text{Aggregation}}$ **Output** $\xrightarrow{\text{Evaluation}}$ **Metrics**

Outcome $\uparrow$ **Metrics**

Lyon and Pacuit 2013

Outcome

Wisdom of crowds



Wisdom of crowds: Sport Obermeyer

Candidates → Recruitment → **Crowd** → Elicitation → **Inputs** → Aggregation → **Output** → Evaluation → **Metrics**

Inputs

30	150	80
90	40	50
300	80	170

Output

Normal

μ	140	90	100
σ	141	56	62

$$Q^* = \mu + \sigma\Phi^{-1}\left(\frac{c_u}{c_u + c_o}\right)$$

Metrics

MAPE: 56%

NV profit: $20K

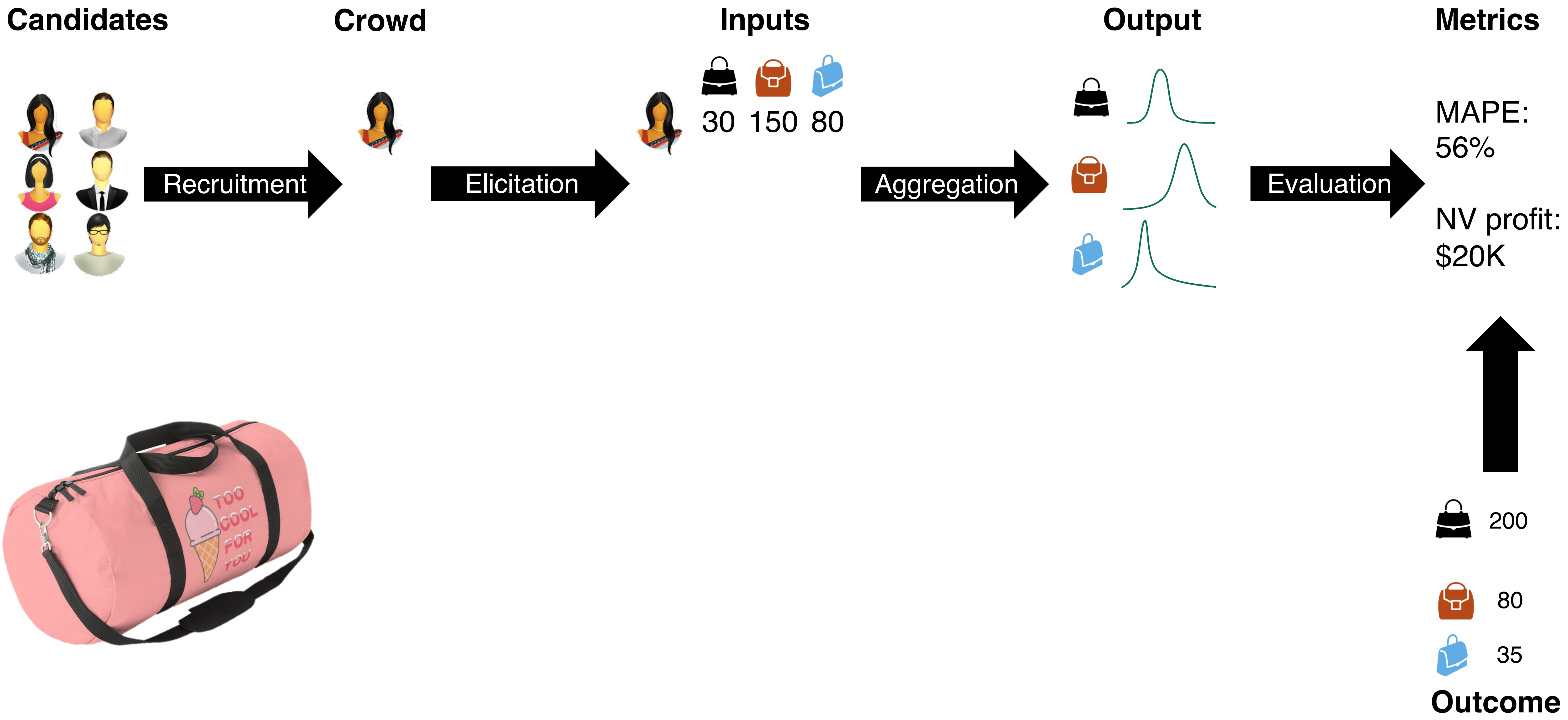
200

80

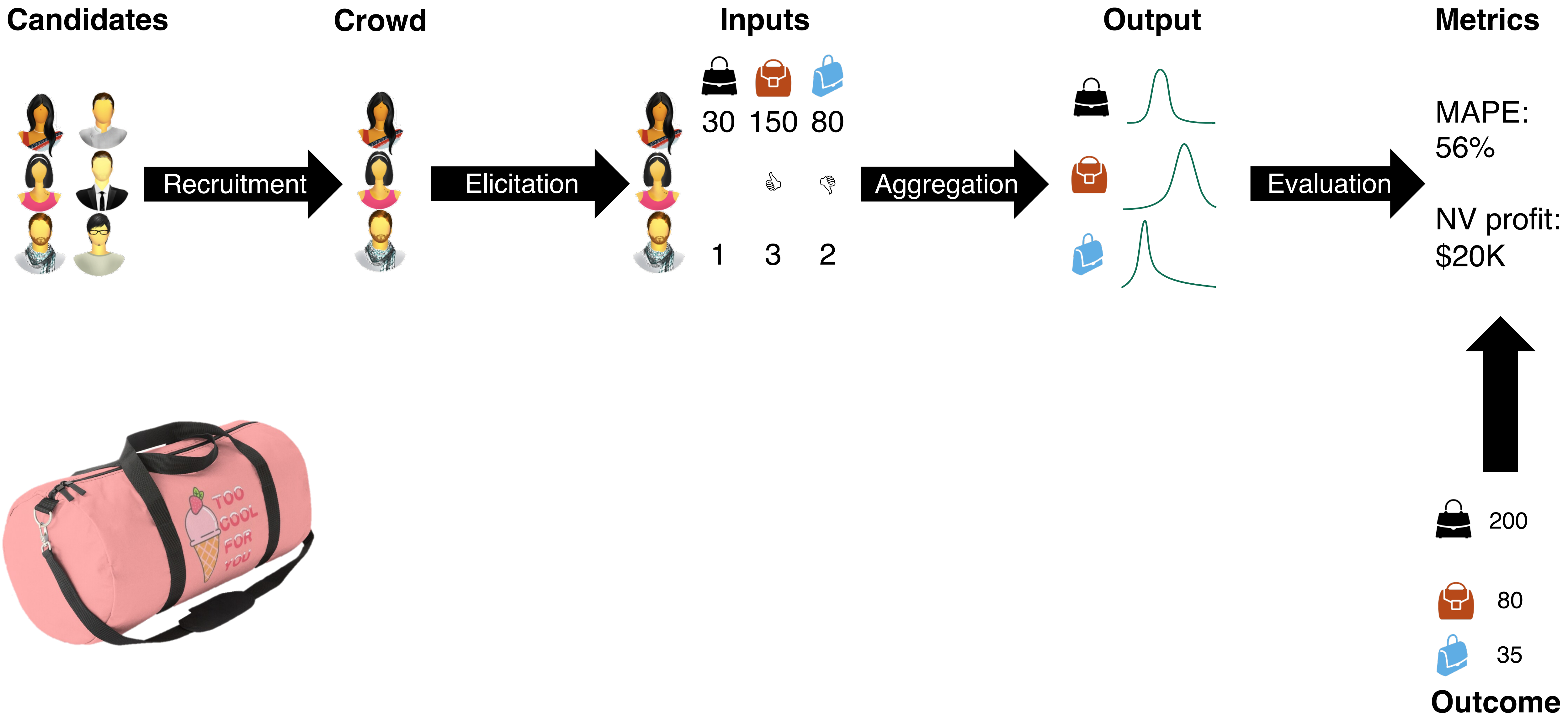
35

Outcome

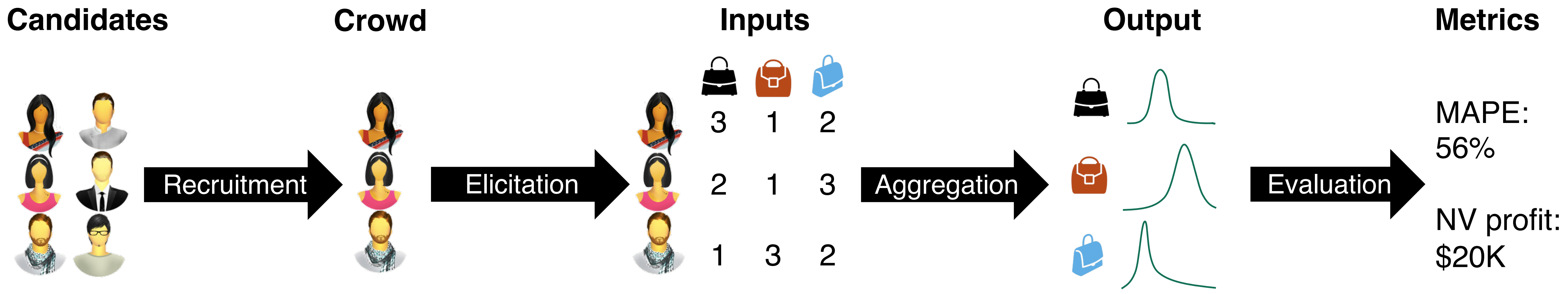
Wisdom of crowds: challenges

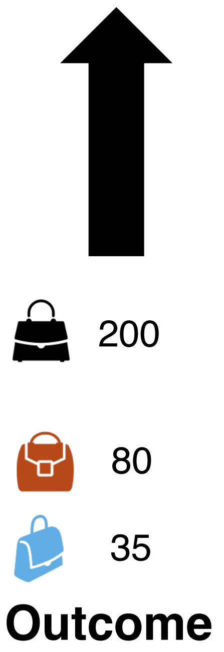


Wisdom of crowds: challenges

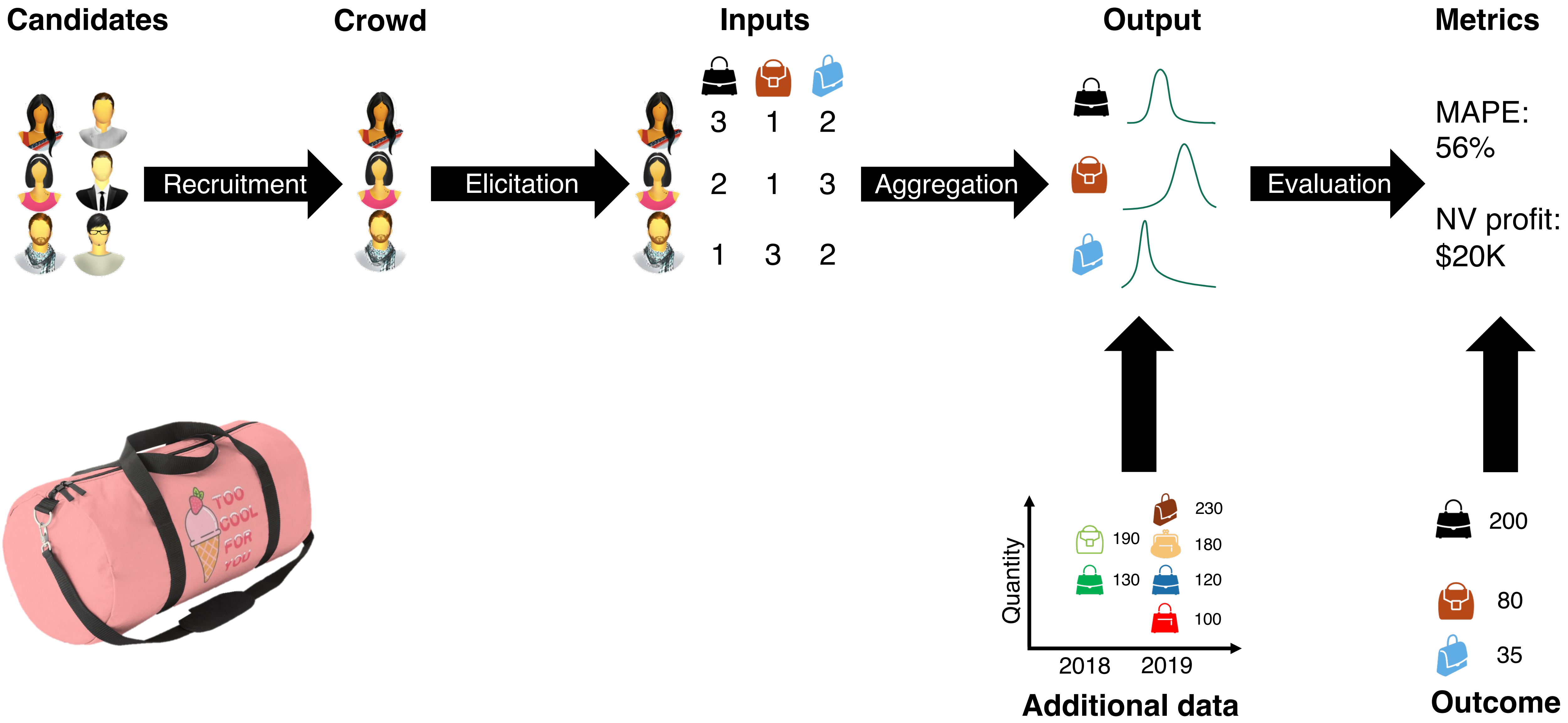


Wisdom of crowds: challenges

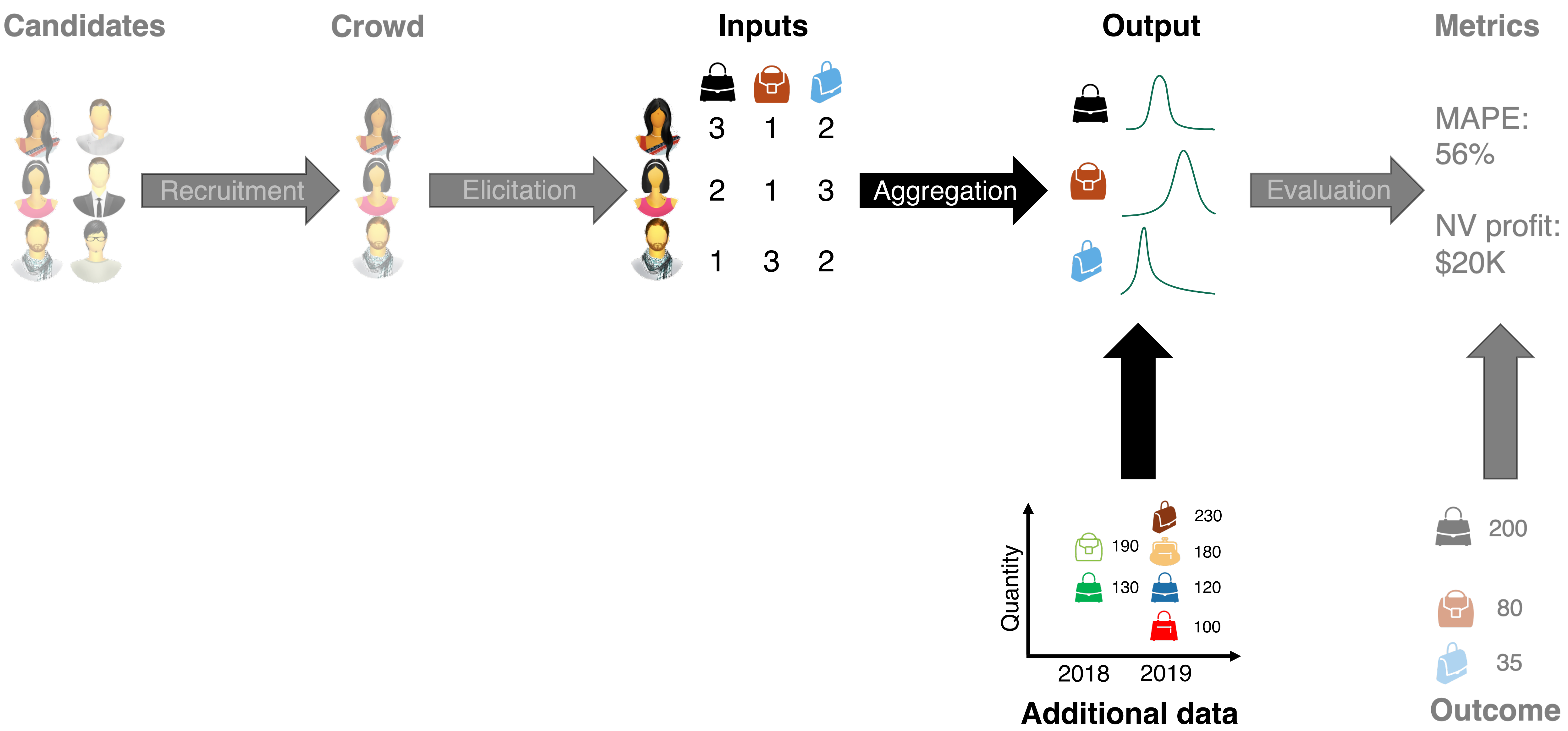




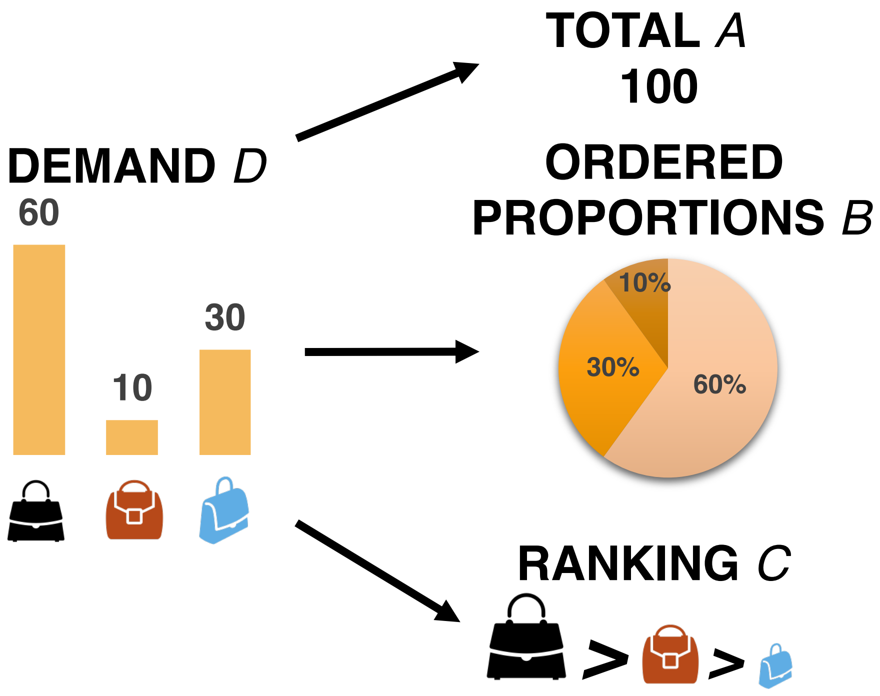
Wisdom of crowds: challenges



Aggregation



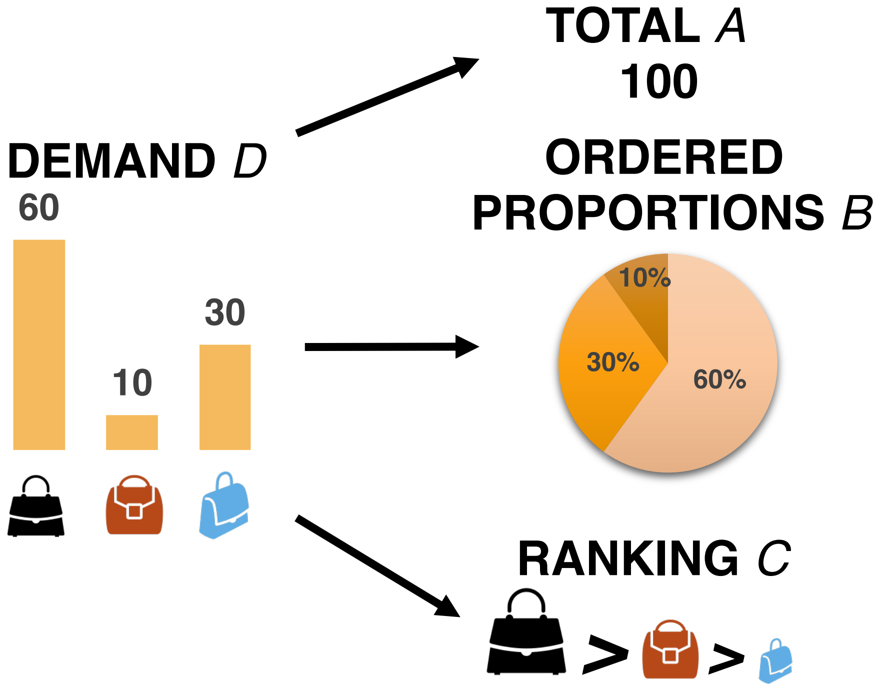
Decomposition



Matrix representation

$$\overbrace{[60, 10, 30]}^{D} = \overbrace{100}^{A} \times \overbrace{[0.6, 0.3, 0.1]}^{B} \times \overbrace{\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{bmatrix}}^{C}$$

Decomposition



Matrix representation

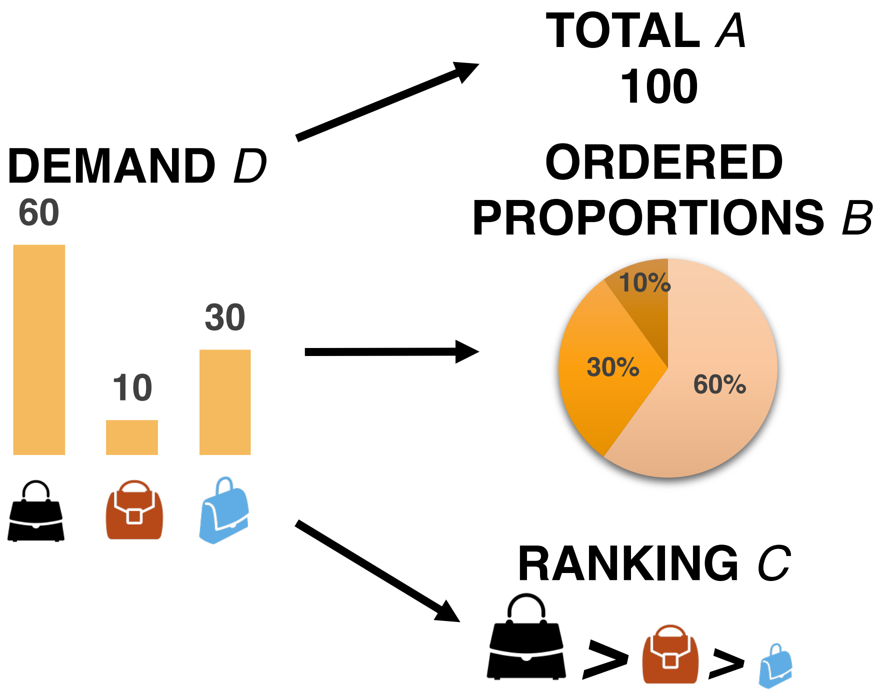
$$\overbrace{[60, 10, 30]}^{D} = \overbrace{100}^{A} \times \overbrace{[0.6, 0.3, 0.1]}^{B} \times \overbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}}^{C}$$

Forecasts are random variables

m products

$$\overbrace{\mathcal{D}}^{1\times m} = \overbrace{\mathcal{A}}^{1\times 1} \times \overbrace{\mathcal{B}}^{1\times m} \times \overbrace{\mathcal{C}}^{m\times m}$$

Decomposition



Matrix representation

$$\overbrace{[60, 10, 30]}^{D} = \overbrace{100}^{A} \times \overbrace{[0.6, 0.3, 0.1]}^{B} \times \overbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}}^{C}$$

Forecasts are random variables

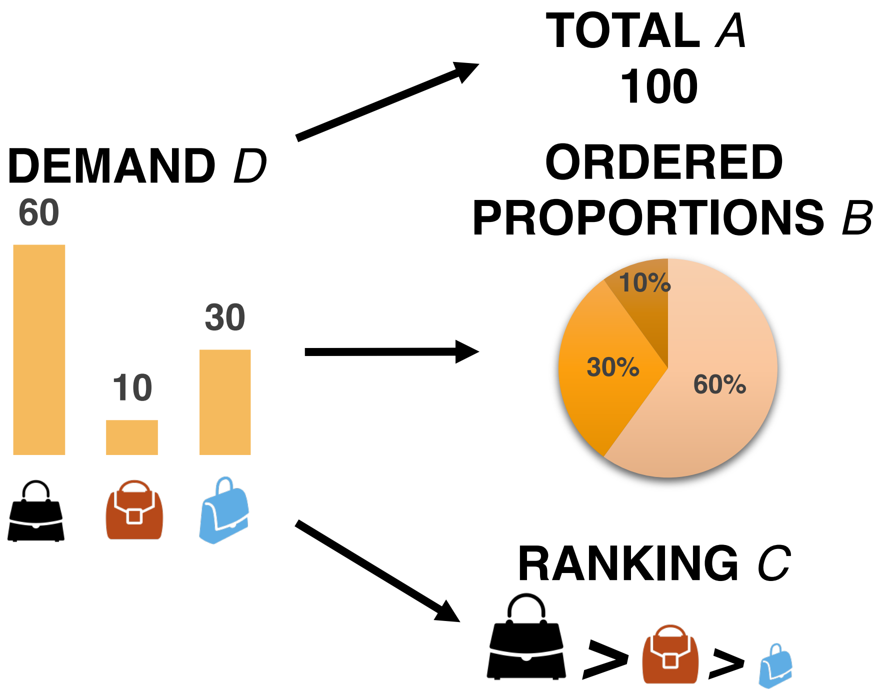
m products

$$\overbrace{\mathcal{D}}^{1\times m} = \overbrace{\mathcal{A}}^{1\times 1} \times \overbrace{\mathcal{B}}^{1\times m} \times \overbrace{\mathcal{C}}^{m\times m}$$

Scenario

	ω_1	ω_2
Probability	0.5	0.5
$\mathcal{D}_1$	400	150
$\mathcal{D}_2$	600	800
$\mathcal{D}_3$	1000	50
Total $\mathcal{A}$	2000	400
Top-1 prop. $\mathcal{B}_1$	0.5	0.8
Top-2 prop. $\mathcal{B}_2$	0.3	0.15
Ranking	$\mathcal{D}_3 > \mathcal{D}_2 > \mathcal{D}_1$	$\mathcal{D}_2 > \mathcal{D}_3 > \mathcal{D}_1$

Decomposition



Matrix representation

$$\overbrace{[60, 10, 30]}^{D} = \overbrace{100}^{A} \times \overbrace{[0.6, 0.3, 0.1]}^{B} \times \overbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}}^{C}$$

Forecasts are random variables

m products

$$\overbrace{\mathcal{D}}^{1 \times m} = \overbrace{\mathcal{A}}^{1 \times 1} \times \overbrace{\mathcal{B}}^{1 \times m} \times \overbrace{\mathcal{C}}^{m \times m}$$

Scenario

	ω_1	ω_2
Probability	0.5	0.5
$\mathcal{D}_1$	400	150
$\mathcal{D}_2$	600	800
$\mathcal{D}_3$	1000	50
Total $\mathcal{A}$	2000	400
Top-1 prop. $\mathcal{B}_1$	0.5	0.8
Top-2 prop. $\mathcal{B}_2$	0.3	0.15
Ranking	$\mathcal{D}_3 > \mathcal{D}_2 > \mathcal{D}_1$	$\mathcal{D}_2 > \mathcal{D}_3 > \mathcal{D}_1$

We can also rebuild forecasts from components

Recombination by simulation

Total $\mathcal{A}$	**Ordered proportions $\mathcal{B}$**	**Ranking $\mathcal{C}$**

Recombination by simulation

Total $\mathcal{A}$

100

Ordered proportions $\mathcal{B}$

1st	**2nd**	**3rd**
0.8	0.15	0.05

Ranking $\mathcal{C}$

1 3 2

Recombination by simulation

Total $\mathcal{A}$

100

Ordered proportions $\mathcal{B}$

1st	**2nd**	**3rd**
0.8	0.15	0.05

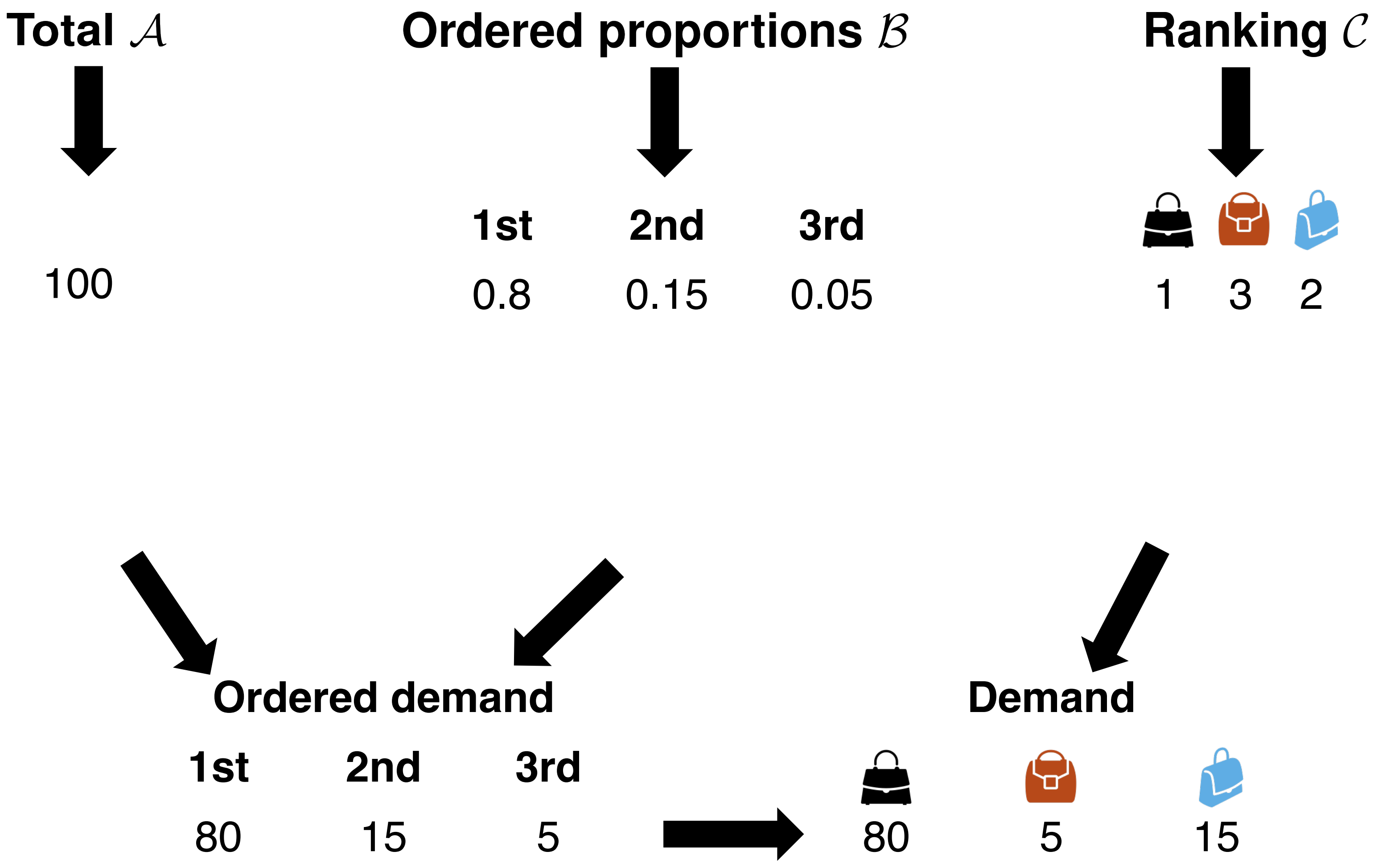
Ranking $\mathcal{C}$

1 3 2

Ordered demand

1st	**2nd**	**3rd**
80	15	5

Recombination by simulation



Recombination by simulation

Total $\mathcal{A}$

100
120
100
150

Ordered proportions $\mathcal{B}$

1st	2nd	3rd
0.8	0.15	0.05
0.6	0.3	0.1
0.75	0.15	0.1
0.5	0.3	0.2

Ranking $\mathcal{C}$

1	3	2
1	3	2
2	3	1
3	1	2

Ordered demand

1st	2nd	3rd
80	15	5
72	36	12
75	15	10
50	30	20

Demand

80	5	15
72	12	36
15	10	75
30	75	45

Recombination by simulation

Total $\mathcal{A}$ **Ordered proportions $\mathcal{B}$** **Ranking $\mathcal{C}$**

100
120
100
150

1st	2nd	3rd
0.8	0.15	0.05
0.6	0.3	0.1
0.75	0.15	0.1
0.5	0.3	0.2

1	3	2
1	3	2
2	3	1
3	1	2

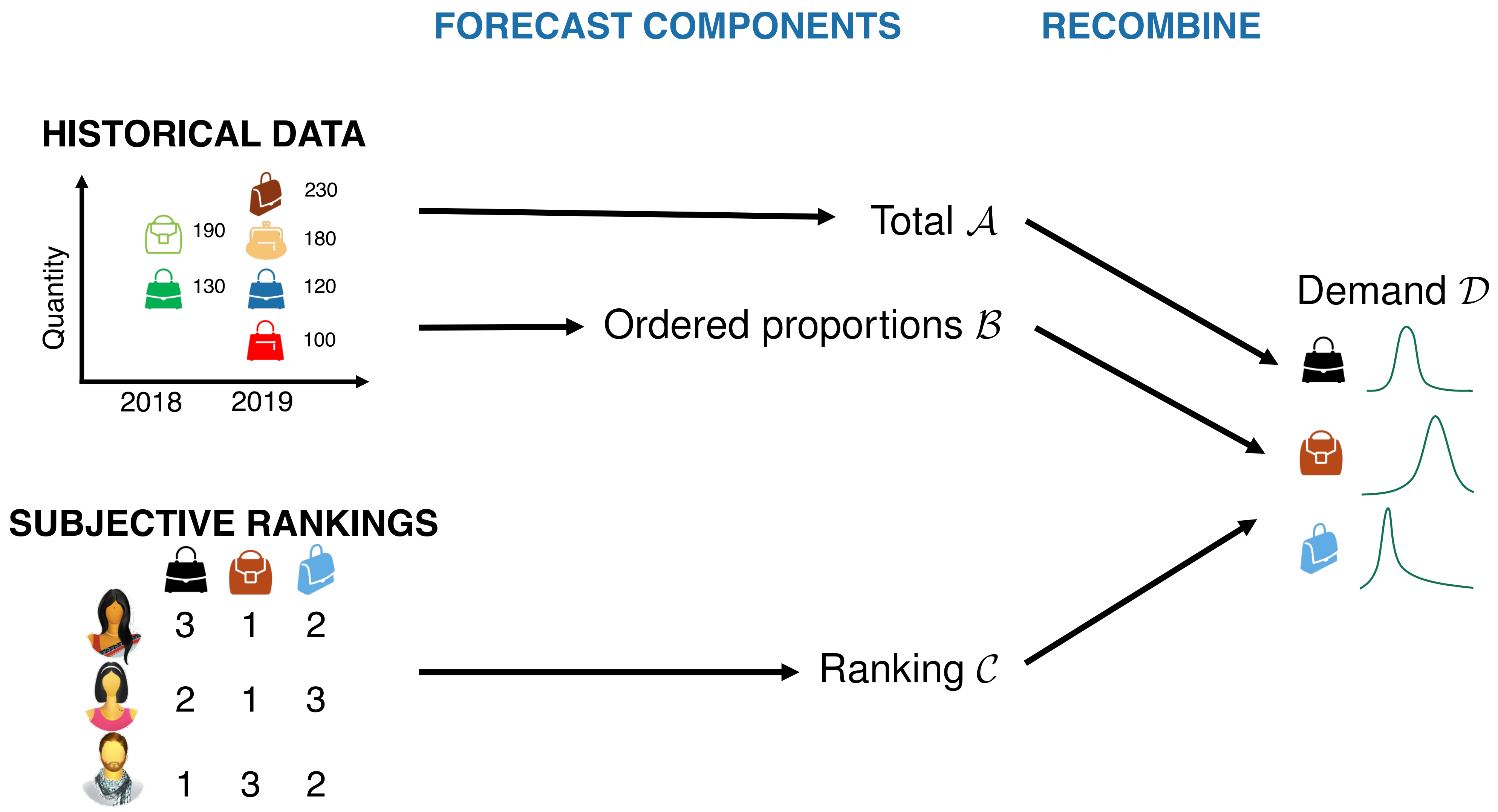
Ordered demand

1st	2nd	3rd
80	15	5
72	36	12
75	15	10
50	30	20

Demand

80	5	15
72	12	36
15	10	75
30	75	45

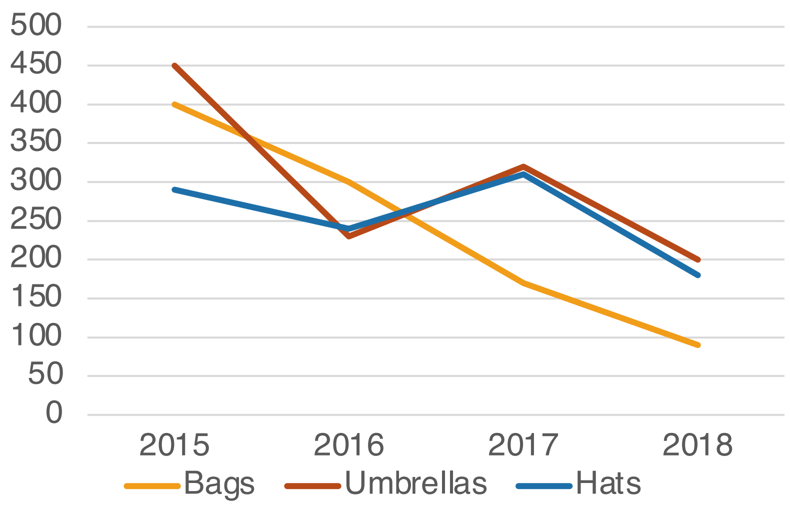
Component models



Total demand

DEMAND D = **TOTAL A** x ORDERED PROPORTIONS B x RANKING C

Historical data



Often, only 2 or 3 seasons per category.

Solution: pool data from multiple categories.

Regression models

Linear models based on category-level covariates

Examples: category-level fixed effects, # of products, time trend

Specifications: Gaussian model, Gamma-GLM

Ordered proportions

DEMAND *D* = TOTAL *A* x ORDERED PROPORTIONS *B* x RANKING *C*

Historical data

		Rank				
		1st	**2nd**	**3rd**	**4th**	**5th**
	2016	0.3	0.25	0.2	0.15	0.1
Season	2017	0.6	0.3	0.1		
	2018	0.4	0.3	0.2	0.1	

Different number of products in each season

Ordered proportions

DEMAND D = **TOTAL** A x **ORDERED PROPORTIONS** B x **RANKING** C

Example of a model

Gamma

$u_i \sim \text{Gamma}(\cdot \mid 1, \lambda_B)$

$+$

MNL choice model

$$p_i = \frac{u_i}{\sum_i u_i}$$

$\longrightarrow$ $\text{Dirichlet}(\lambda_B, \ldots, \lambda_B)$

Ordered proportions

DEMAND D = TOTAL A x **ORDERED PROPORTIONS B** x RANKING C

General model

Utilities → **Choice model** → **Proportions** → **Sort** → **Ord. prop.**

Utilities			Proportions			Ord. prop.		
						1st	**2nd**	**3rd**
4	10	6	0.2	0.5	0.3	0.5	0.3	0.2

$u_i \sim F(\cdot \mid \lambda_B)$ $\quad\quad$ $\mathbf{p} = G(\mathbf{u})$ $\quad\quad$ $\mathbf{B} = \mathrm{Sort}(\mathbf{p})$

Example of a model

Gamma

$u_i \sim \mathrm{Gamma}(\cdot \mid 1, \lambda_B)$

$+$

MNL choice model

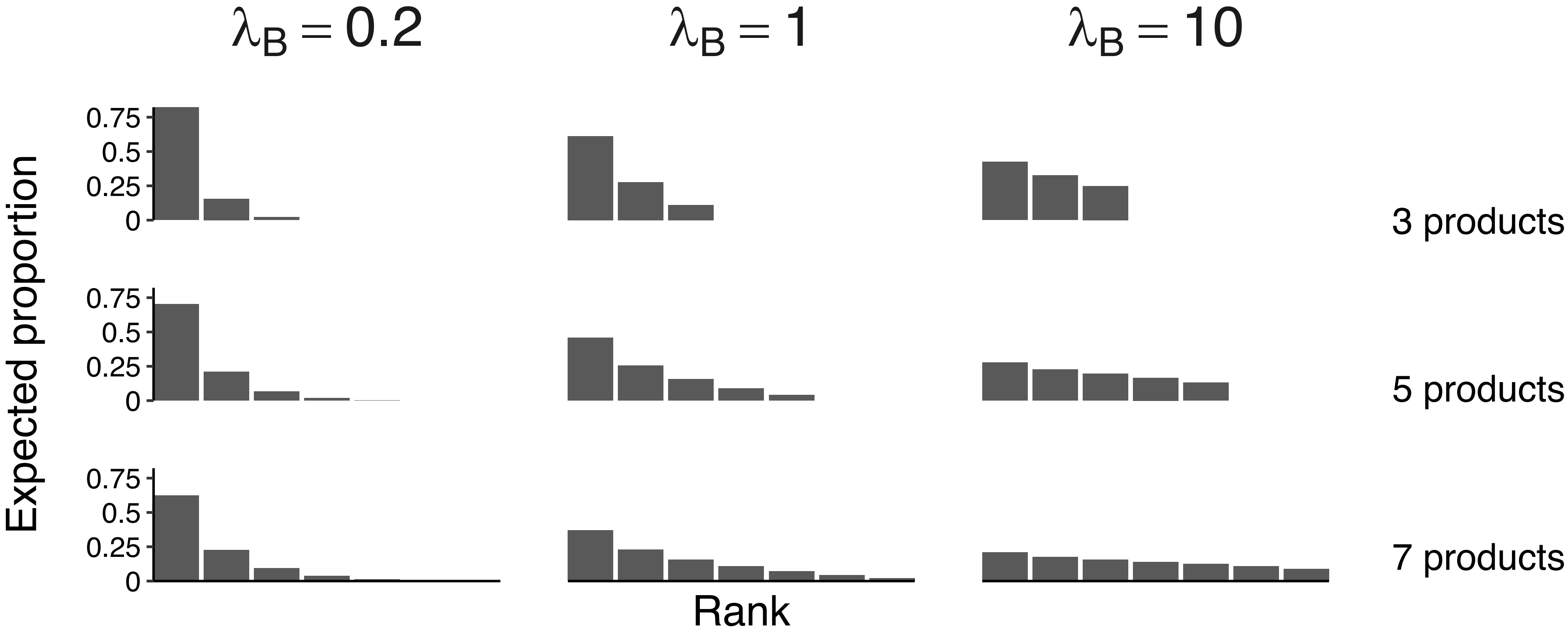
$$p_i = \frac{u_i}{\sum_i u_i} \longrightarrow \mathrm{Dirichlet}(\lambda_B, \ldots, \lambda_B)$$

Ordered proportions

DEMAND *D* = **TOTAL** *A x* **ORDERED PROPORTIONS** *B x* **RANKING** *C*

Dirichlet model: higher $\lambda_B \rightarrow$ less concentration

Expected values of the ordered proportion vector for 3, 5 and 7 products



Ranking

DEMAND *D* = TOTAL *A* x ORDERED PROPORTIONS *B* x **RANKING *C***

More disagreement → more uncertainty

Past performance is not available

Subjective inputs

			
	2	1	3
	2	3	1
	1	3	2

Method

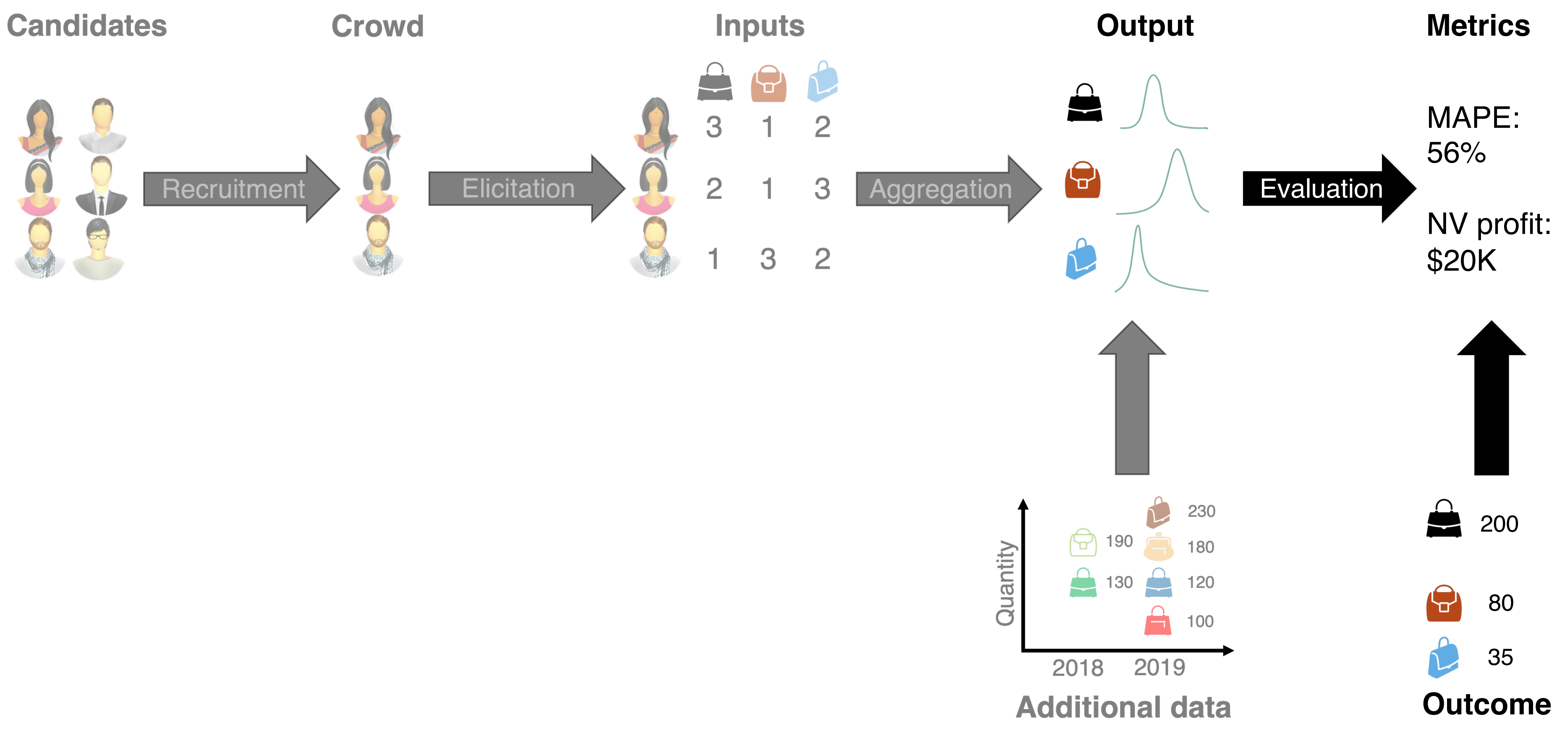
Empirical distribution:

1/3 on 

1/3 on 

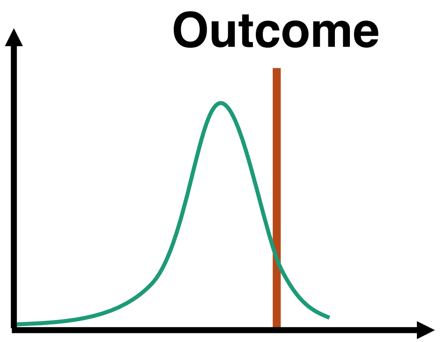
1/3 on 

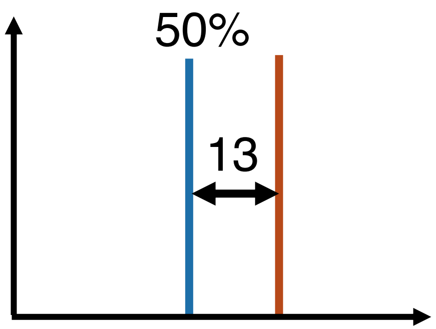
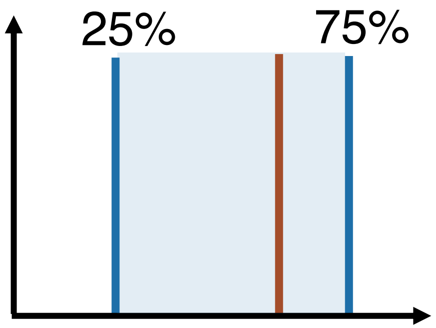
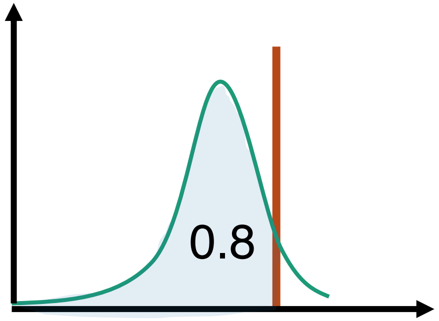
Evaluation



Candidates
Crowd
Inputs
Output
Metrics
Recruitment
Elicitation
Aggregation
Evaluation
3 1 2
2 1 3
1 3 2
MAPE: 56%
NV profit: $20K
Quantity
190
130
230
180
120
100
2018
2019
Additional data
200
80
35
Outcome

Evaluation: distributions over quantities



	SINGLE PRODUCT	MULTIPLE PRODUCTS	BEST CASE
POINT FORECAST		13, 7, 5, 15, 18, 19, 2, 1, 10, 10 Avg. error: 10 units	Avg. error: 0
INTERVAL		1, 0, 1, 0, 0, 0, 1, 0, 0, 0 Avg. coverage: 30%	Avg. coverage: 50%
PROBABILITY INTEGRAL TRANSFORM		0.8, 0.7, 0.9, 0.6, 0.1, 0.8, 0.9, 0.8, 0.5, 0.7	Should be distributed as U(0, 1)

Evaluation: decomposing forecasts

Demand

80	15	5
72	12	36
30	20	150
30	75	45

Total
100
120
200
150

LEVEL

1st	**2nd**	**3rd**
0.8	0.15	0.05
0.6	0.3	0.1
0.75	0.15	0.1
0.5	0.3	0.2

CONCENTRATION

1	2	3
1	3	2
2	3	1
3	1	2

RESOLUTION

Evaluation: decomposing forecasts

Demand

80	15	5
72	12	36
30	20	150
30	75	45

Total	**1st**	**2nd**	**3rd**			
100	0.8	0.15	0.05	1	2	3
120	0.6	0.3	0.1	1	3	2
200	0.75	0.15	0.1	2	3	1
150	0.5	0.3	0.2	3	1	2
LEVEL	**CONCENTRATION**			**RESOLUTION**		

Evaluation: scoring rules for rankings

Binary events

Brier score

$$(p - o)^2$$

p — forecast

o — outcome

Proper scoring rule

A forecaster reports honestly if rewarded according to Brier score

Interpretable scale

0 – the best value
1 – the worst value
0.25 – uniform distribution

Evaluation: scoring rules for rankings

Binary events

Brier score

$$(p - o)^2$$

p — forecast

o — outcome

Proper scoring rule

A forecaster reports honestly if rewarded according to Brier score

Interpretable scale

0 – the best value
1 – the worst value
0.25 – uniform distribution

Rankings

Brier score for distributions over rankings?

Should also punish **slight** mistakes less than **large** mistakes

(Example: true outcome 1 > 2 > 3, then 3 > 2 > 1 is worse than 1 > 3 > 2).

Solution: apply Gneiting and Raftery (2007) *kernel scoring rule*

$$\underbrace{\mathbb{E}\mathrm{d}(\mathcal{C}, \mathrm{C})}_{\text{avg. error}} - \frac{1}{2}\underbrace{\mathbb{E}\mathrm{d}(\mathcal{C}, \mathcal{C}')}_{\text{dispersion}}$$

$\mathcal{C}, \mathcal{C}'$ — indep. copies of ranking forecast

C — actual ranking

$\mathrm{d}(\cdot, \cdot)$ — **kernel**, essentially a distance

Evaluation: scoring rules for rankings

Binary events

Brier score

$$(p - o)^2$$

p — forecast

o — outcome

Proper scoring rule

A forecaster reports honestly if rewarded according to Brier score

Interpretable scale

0 – the best value
1 – the worst value
0.25 – uniform distribution

Rankings

Brier score for distributions over rankings?

Should also punish **slight** mistakes less than **large** mistakes

(Example: true outcome 1 > 2 > 3, then 3 > 2 > 1 is worse than 1 > 3 > 2).

Solution: apply Gneiting and Raftery (2007) *kernel scoring rule*

$$\underbrace{\mathbb{E}\mathrm{d}(\mathcal{C}, \mathrm{C})}_{\text{avg. error}} - \frac{1}{2} \underbrace{\mathbb{E}\mathrm{d}(\mathcal{C}, \mathcal{C}')}_{\text{dispersion}}$$

$\mathcal{C}, \mathcal{C}'$ — indep. copies of ranking forecast

C — actual ranking

$\mathrm{d}(\cdot, \cdot)$ — **kernel**, essentially a distance

Kernels for rankings

Map rankings to vectors

Vectors should be rescaled

Compute Euclidean distance between these vectors

Example:

$1 > 2 > 3 \to [1, 2, 3]/\sqrt{8}$

$3 > 2 > 1 \to [3, 2, 1]/\sqrt{8}$

$d(1 > 2 > 3, 3 > 2 > 1) = 1$

because

$1/8(4 + 0 + 4) = 1$

Scoring rule corresponding to this kernel: I call it **Spearman-Brier** score.

Empirical results

moods of norway

Dataset

2 seasons of historical data

253 products, 52 categories

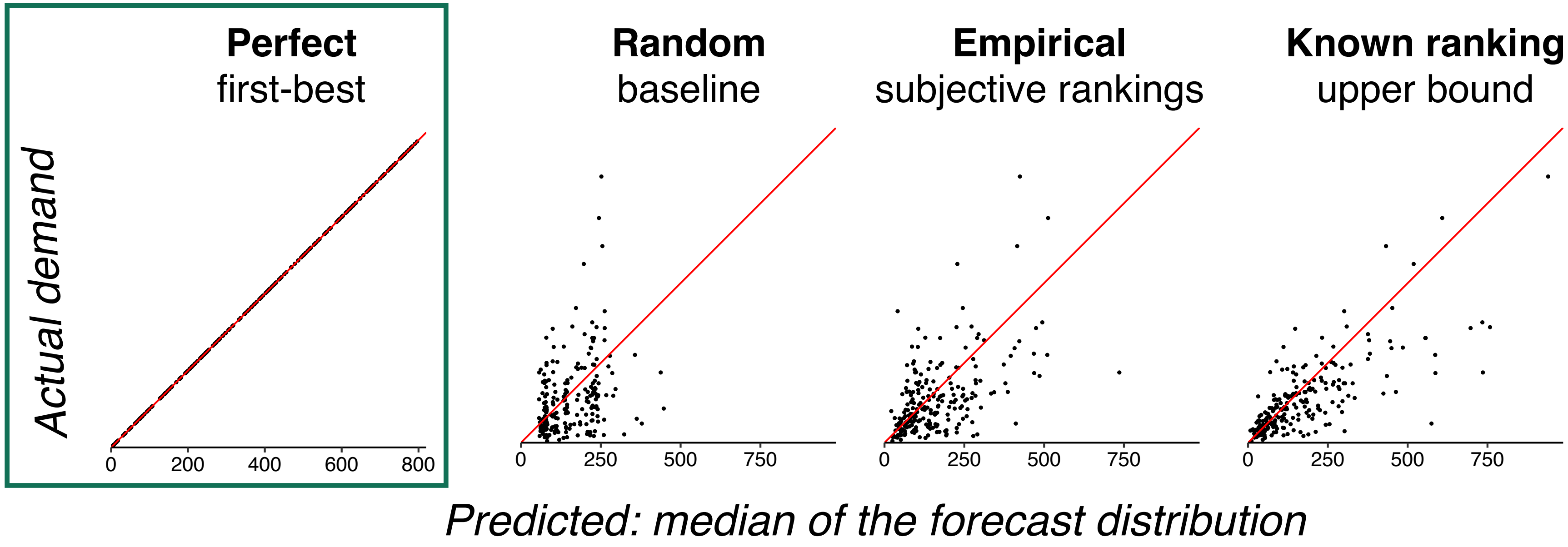
2460 forecasts by 21 experts

Models

Total: log-normal regression

Ord. prop: Dirichlet

Ranking: empirical



Empirical results

Dataset

2 seasons of historical data

253 products, 52 categories

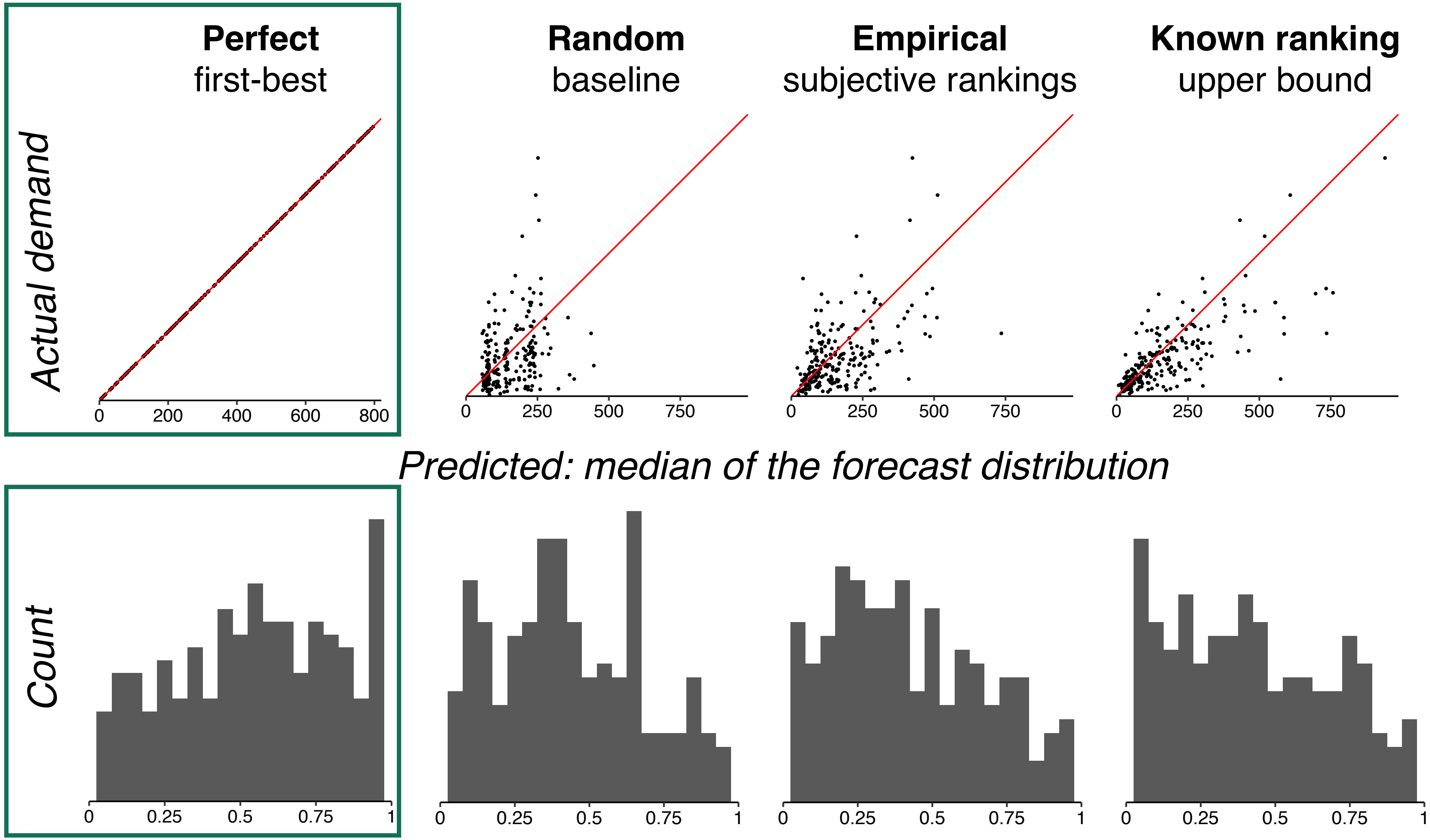
2460 forecasts by 21 experts

Models

Total: log-normal regression

Ord. prop: Dirichlet

Ranking: empirical



Predicted: median of the forecast distribution

PIT: forecast CDF evaluated at the outcome

Empirical results: components

Dataset
2 seasons of historical data

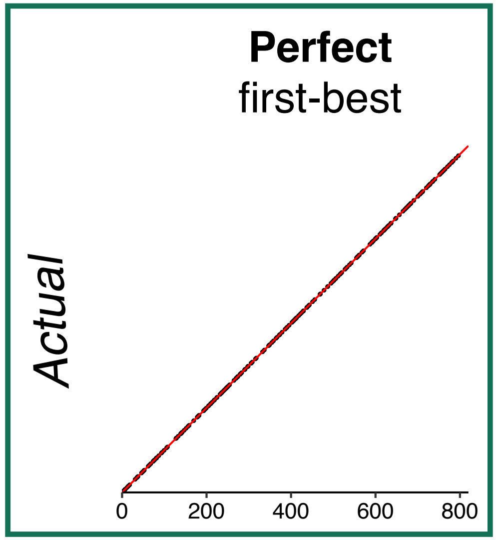
253 products, 52 categories

2460 forecasts by 21 experts

Models
Total: log-normal regression

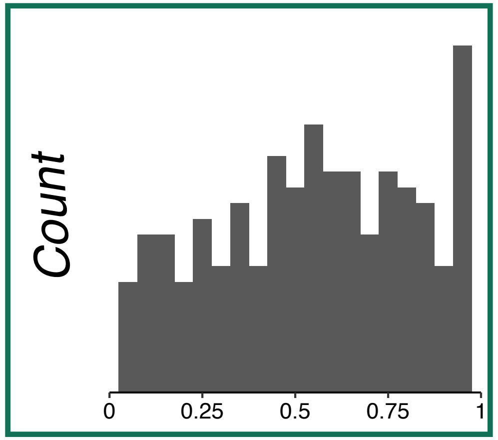
Ord. prop: Dirichlet

Ranking: empirical



Total | **Ordered proportions** | **Ranking**

Predicted: median of the forecast distribution



PIT: forecast CDF evaluated at the outcome

Empirical results: components

moods of norway

Dataset

2 seasons of historical data

253 products, 52 categories

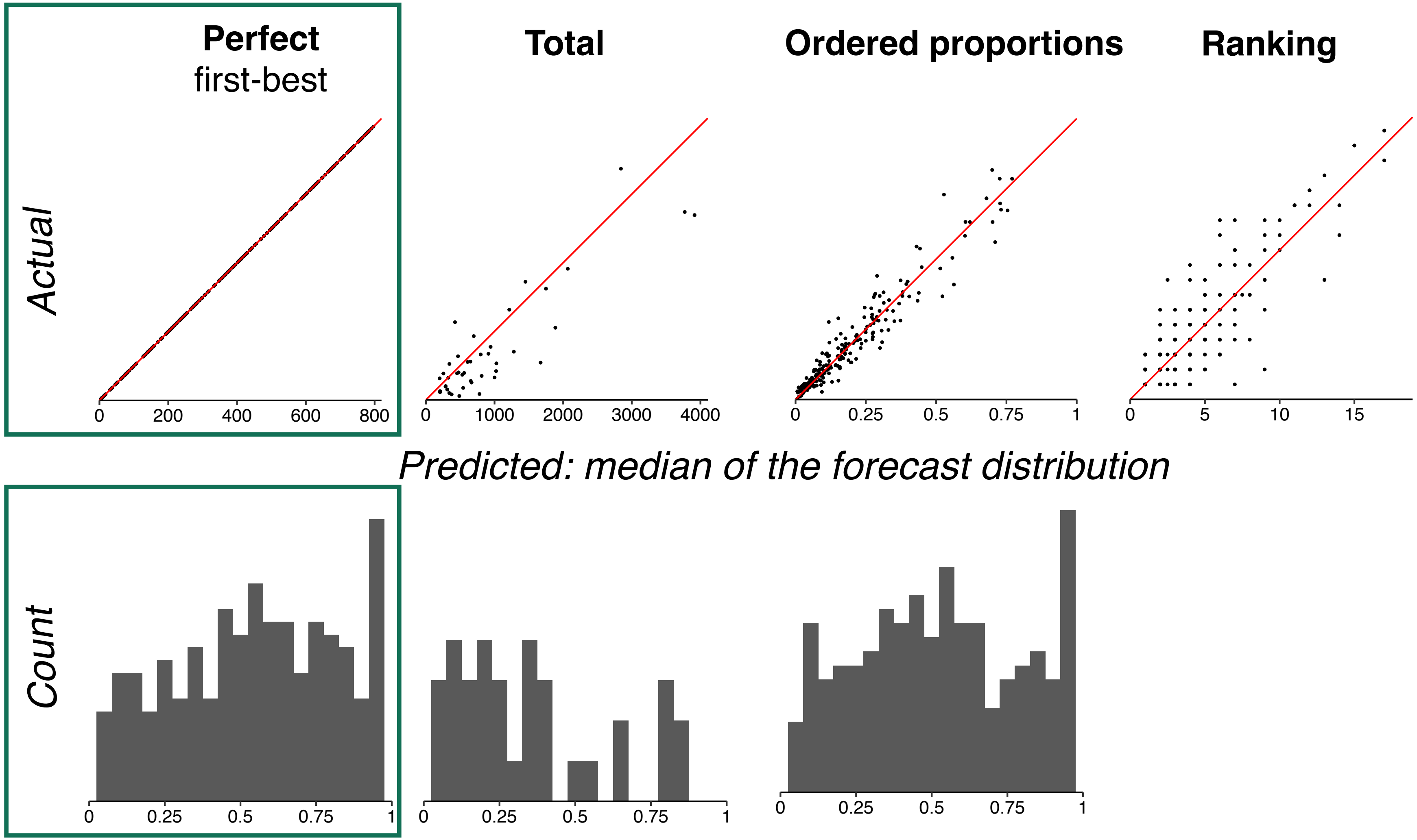
2460 forecasts by 21 experts

Models

Total: log-normal regression

Ord. prop: Dirichlet

Ranking: empirical



Predicted: median of the forecast distribution

PIT: forecast CDF evaluated at the outcome

Empirical results: metrics

moods of norway

Dataset

2 seasons of historical data

253 products, 52 categories

2460 forecasts by 21 experts

Models

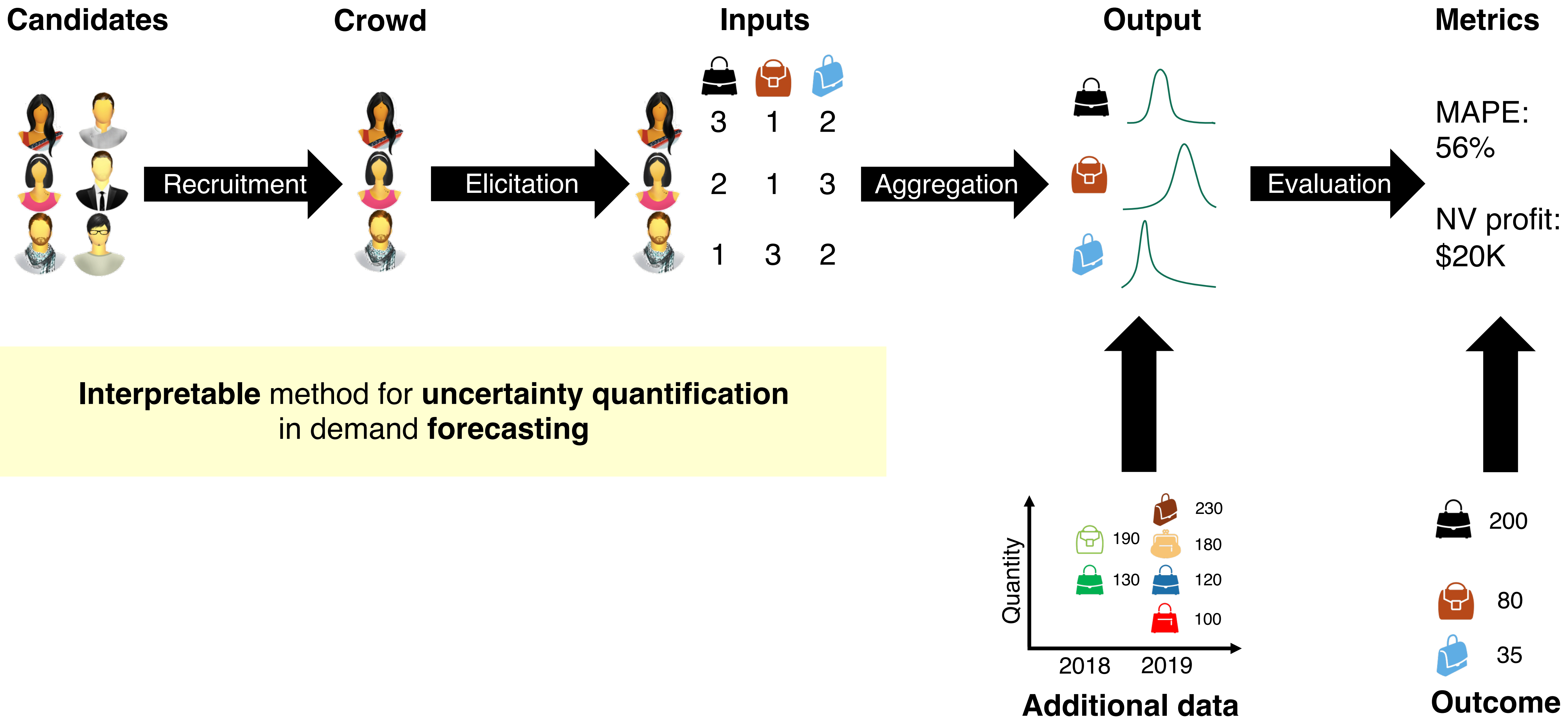
Total: log-normal regression

Ord. prop: Dirichlet

Ranking: empirical

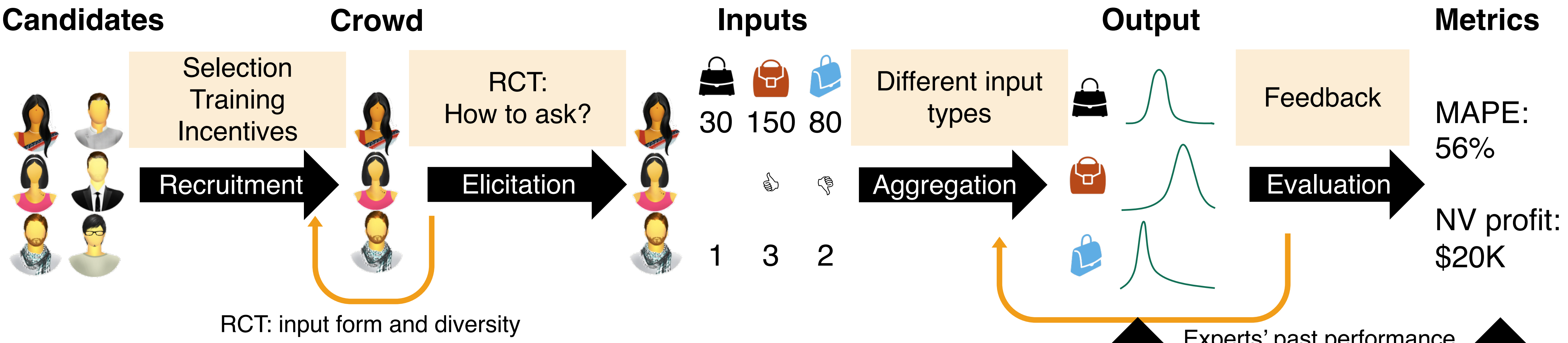
Metric	Perfect	Random	Empirical	Known
50% coverage	0.5	0.62	0.65	0.62
95% coverage	0.95	0.99	0.99	0.97
MAPE	0%	103%	82%	64%
Spearman-Brier score	0	0.25	0.20	0
% of perfect NV profit	100%	59%	61%	65%

Future directions

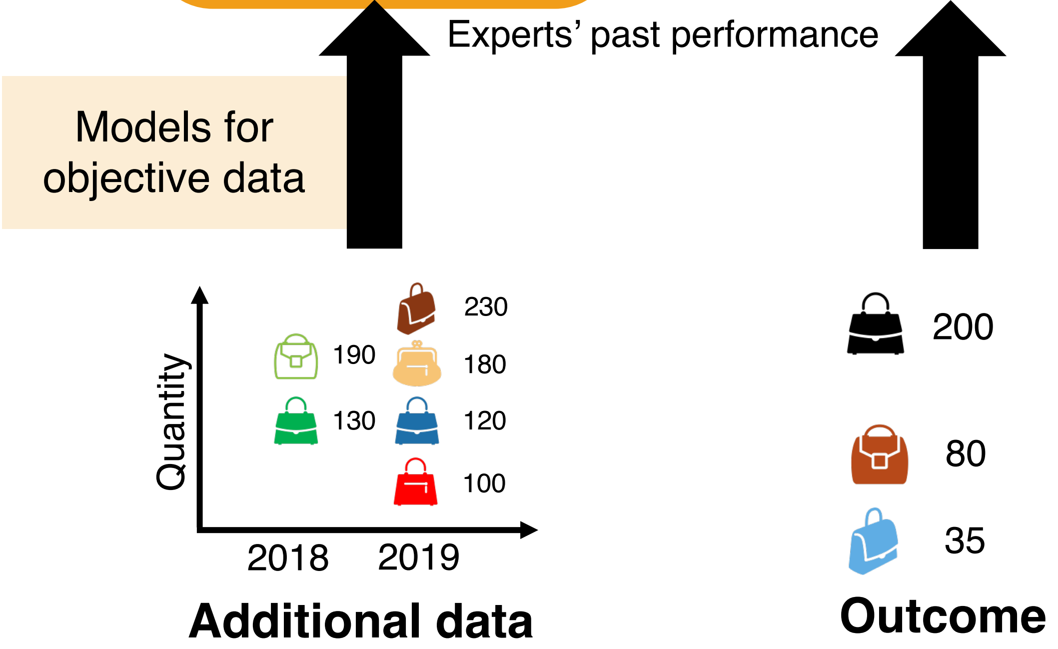


Interpretable method for **uncertainty quantification** in demand **forecasting**

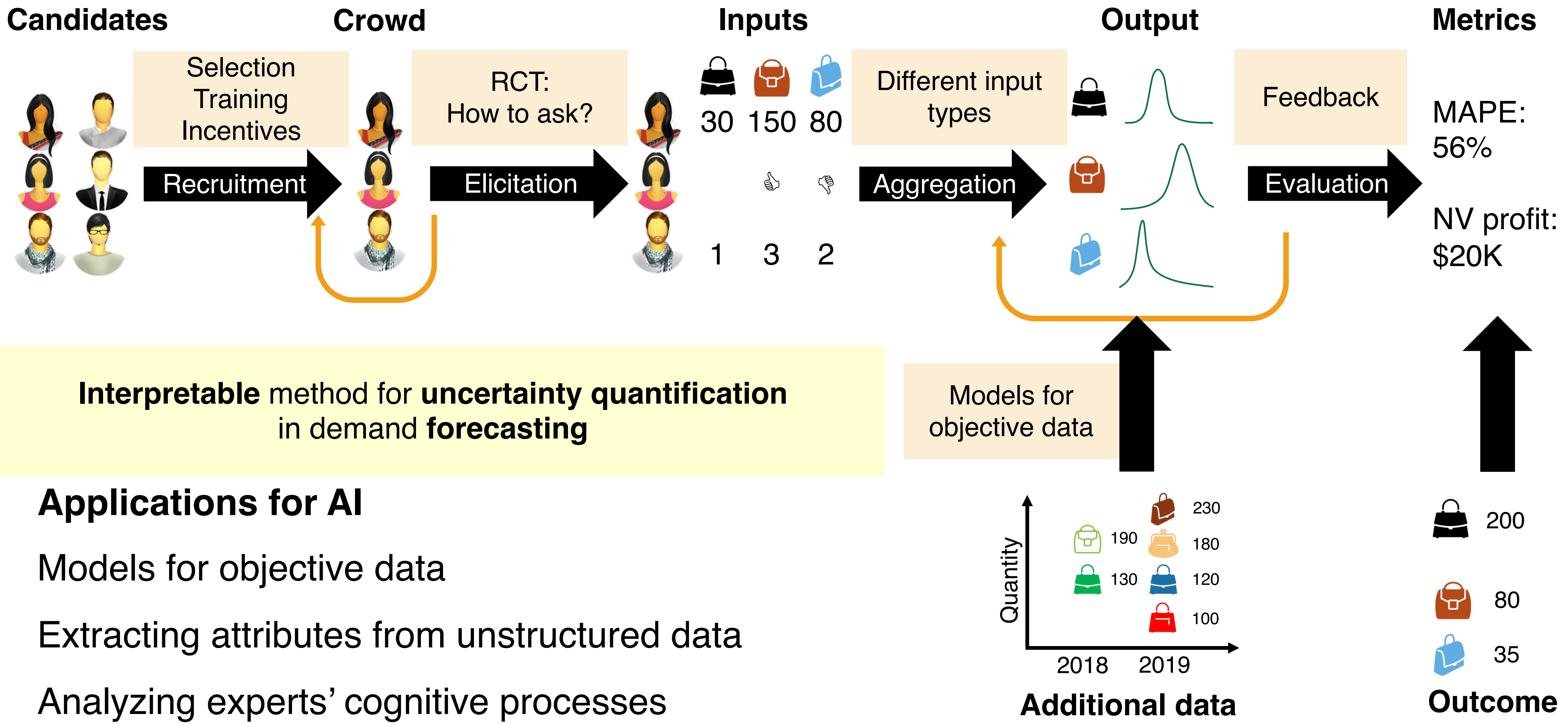
Future directions



Interpretable method for **uncertainty quantification** in demand **forecasting**



Future directions



Interpretable method for **uncertainty quantification** in demand **forecasting**

Applications for AI

Models for objective data

Extracting attributes from unstructured data

Analyzing experts' cognitive processes

BACKUP SLIDES

Research summary

PAPER	CO-AUTHORS	STATUS	MAIN FINDING
1. Forecasting Demand for New Products: Combining Subjective Rankings with Sales Data	Nils Rudi	**major revision,** *Management Science*	We can use rankings from experts to forecast demand.
2. Bias, Information, Noise: The BIN Model of Forecasting	Ville Satopää, Philip Tetlock and Barbara Mellers	**accepted,** *Management Science*	To predict probabilities better, reduce noise.
3. Are Goals Scored Just Before Half Time Worth More? A Soccer Myth Statistically Tested	Henrich R. Greve, Jo Nesbø and Nils Rudi	**published,** *PLOS ONE*	If a home team scores just before halftime, it is more likely to win.
4. Impact of Workforce Flexibility on Customer Satisfaction: Empirical Framework & Evidence from a Cleaning Services Platform	Ekaterina Astashkina and Ruslan Momot	**working paper**	Matching high-ability workers to high-difficulty customers does not improve ratings much.

Decomposition: example 2

	O1	O2	O3	O4	O5	O6	O7	O8
D_1	100	600	100	400	240	250	160	750
D_2	300	400	300	600	160	750	240	250
A	400	1000	400	1000	400	1000	400	1000
B_1	0.75	0.6	0.75	0.6	0.6	0.75	0.6	0.75
D_1 < D_2?	Yes	No	No	Yes	Yes	No	No	Yes

Full specification of the model

Notation

k — category

t — time period

m — number of products

A — total demand

B — proportion

C — ranking

Total demand

$$\log A_{kt} \sim N(\beta_{k,A} + \gamma_{m,A} \log m_{kt}, \sigma_A)$$

$$\beta_k \sim N(\mu_A, \tau_A)$$

Ordered proportions

$$B_{kt} \sim \mathrm{OrderedDirichlet}(\lambda_{k,B} \cdot 1_{m_{kt}})$$

Parameters

$$\beta_{k,A}, \gamma_{m,A}, \lambda_{k,B}, \mu_A, \sigma_A, \tau_A$$

Estimated with MLE

Ranking

Empirical distribution over rankings

Demand forecasting: other methods

Attribute-based methods

Need objectively defined attributes

And stable preferences over these methods

Great when we can have it (Baardman 2018)

But we do not always have the luxury

Prediction markets

Trade “stocks” corresponding to demand ranges.

Might be too complex for companies.

But also promising.

Bassamboo, Cui, Moreno 2015

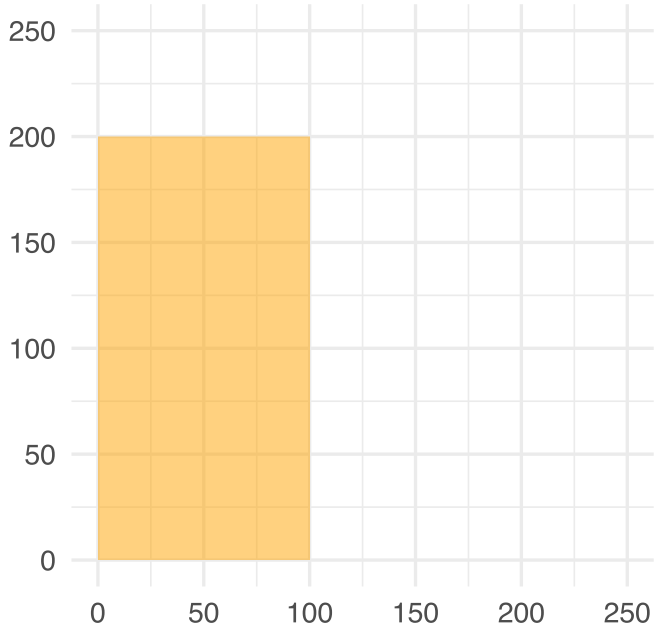
Decomposition: graphical example

$\mathcal{D}_1 \sim U(0, 100)$

$\mathcal{D}_2 \sim U(0, 200)$

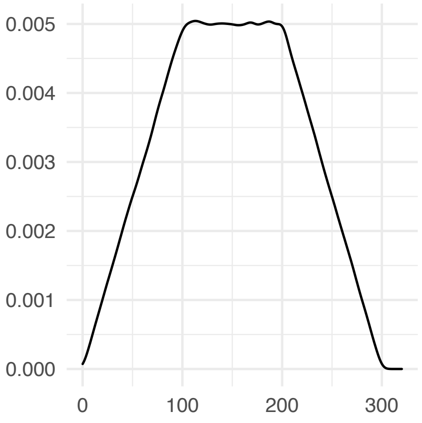
$\mathcal{D}_1$ and $\mathcal{D}_2$ independent

$\mathcal{D}$



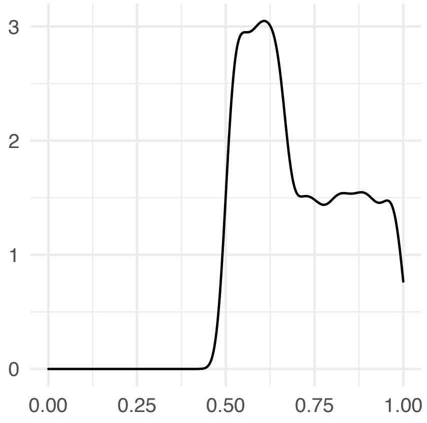
Total

$\mathcal{A}$

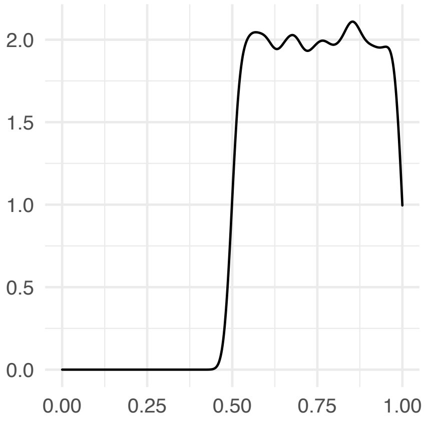


Top-1 proportion given total

$\mathcal{B}_1 \mid \mathcal{A} = 150$



$\mathcal{B}_1 \mid \mathcal{A} = 200$



Ranking given total and ordered proportion

$$\mathbb{P}(\mathcal{D}_1 > \mathcal{D}_2 \mid \mathcal{A} = 150, \mathcal{B}_1 = 0.75) = 0.5$$

$$\mathbb{P}(\mathcal{D}_1 > \mathcal{D}_2 \mid \mathcal{A} = 250, \mathcal{B}_1 = 0.95) = 0$$

Evaluation: implications

	Binary	Quantitative
Calibration	Mean Brier score	Mean squared error
Ranking	ROC-AUC	Spearman-Brier score