

# Mixed Integer Programming Approaches for Real-Time Consumer Preference Elicitation (and Causal Inference)

Juan Pablo Vielma

Massachusetts Institute of Technology

Operations Management/Management Science Workshop  
Booth School of Business, University of Chicago.  
Chicago, IL, October, 2017.

# (Nonlinear) Mixed Integer Programming (MIP)

$$\min \quad f(x)$$

s.t.

$$x \in C$$

$$x_i \in \mathbb{Z} \quad i \in I$$

Mostly **convex**  $f$  and  $C$ .



| Feature         | Chewbacca                           | BB-8                     |
|-----------------|-------------------------------------|--------------------------|
| Wookiee         | Yes                                 | No                       |
| Droid           | No                                  | Yes                      |
| Blaster         | Yes                                 | No                       |
| I would buy toy | <input checked="" type="checkbox"/> | <input type="checkbox"/> |

Marketing And Experimental Design

Causal Inference for Educational Impact of 2010 Chilean Earthquake



“Infinite”-Dimensional MIP and Control or Aerial Drones

# 50+ Years of MIP = Significant Solver Speedups

- Algorithmic Improvements (**Machine Independent**):
  - Commercial Solver Speedup  $\approx 1.9 \times / \text{year}$



- Mostly linear, but also quadratic:
  - Gurobi v6.0 (2014) – v6.5 (2015) quadratic: **4.43 x** (V., Dunning, **Huchette**, Lubin, 2015)
- Also great “open-source” solvers



CBC



GLPK

- Emerging: General Convex Nonlinear (e.g. SDP)



<

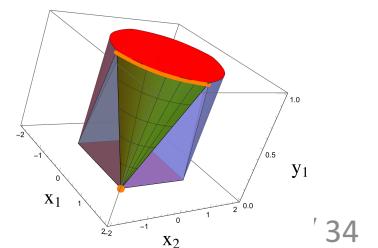
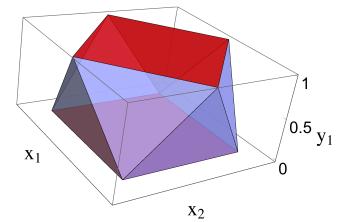
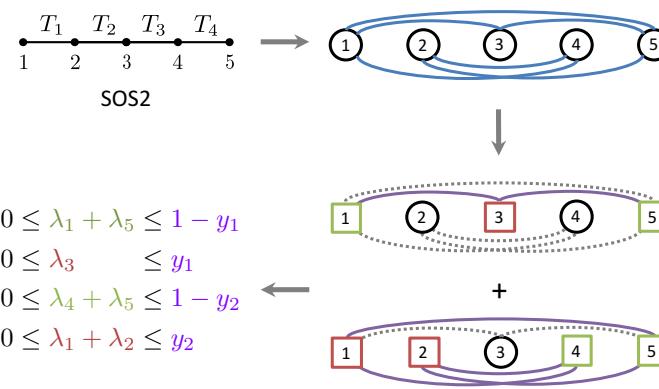
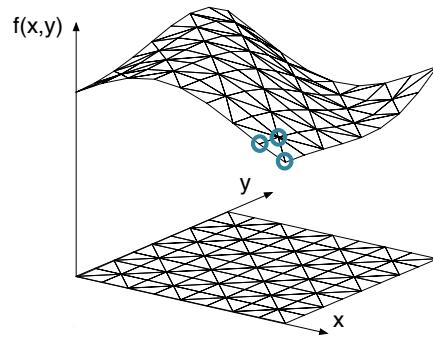


Bonmin <



# MIP Modelling and Advanced Formulations

- MIP Representability: What can be modeled with MIP?
  - Linear: Jeroslow & Lowe '80s ... Basu, Martin, Ryan and Wang '17
  - Convex Nonlinear:
    - MIP formulation for the set of **Prime Numbers**
    - ✓ Non-Convex Polynomial MIP formulation (Jones et al. '76)
    - ✗ Convex of any kind (Lubin, Zadik and V. '17)
- Linear/nonlinear formulation techniques:

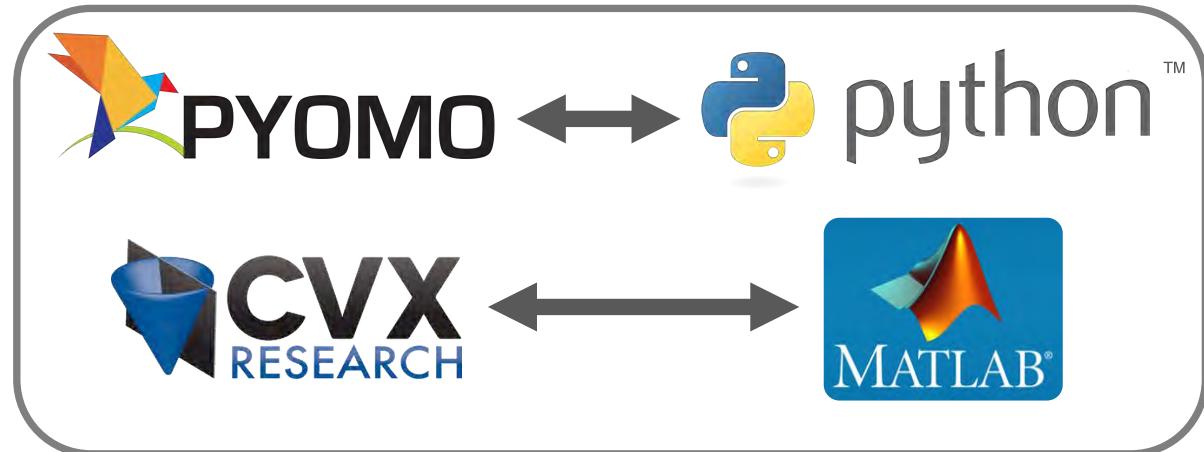


# Accessing Solvers = Modelling Languages

- User-friendly algebraic modelling languages (AML):

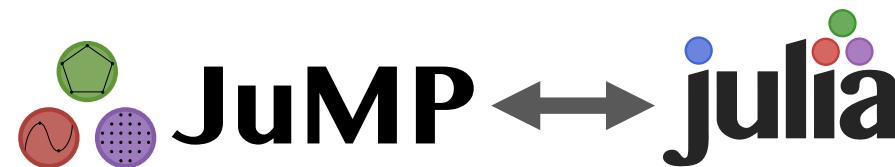


Standalone and Fast

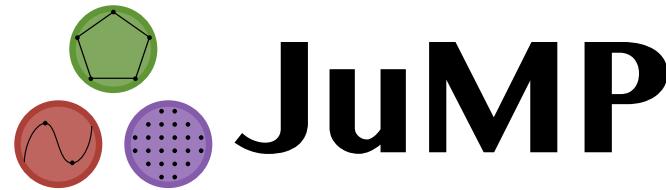


Based on General Language and Versatile

- 21<sup>st</sup> Century AML:



- Free and Open-Source
- Easy to use, but as advanced as proprietary C/C++ interphases
- As fast as standalone AMLs and C/C++ interphases

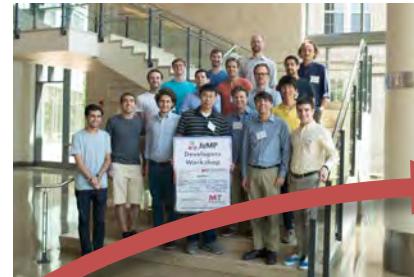


Created by students



Iain Dunning, Miles Lubin  
and **Joey Huchette**  
2016 ICS Prize

Community Developers



Software Engineer



Jarrett  
Revels



JuMP-Suit?



Juan Pablo  
Vielma



# Outline

---

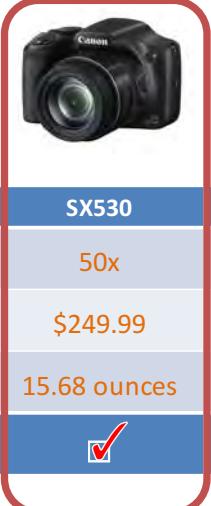
- MIP and Consumer Preference Elicitation
  - Direct improvement from MIP formulation
  - Fast, versatile and efficient learning
- MIP and Causal Inference
  - Indirect improvement from MIP formulation
  - Right formulation brings you back to solving the problem you really wanted to solve

Mixed Integer Programming  
(joint work with Joey Huchette)

and

Consumer Preference Elicitation  
(joint work with Denis Saure)

# Adaptive Choice-Based Conjoint Analysis



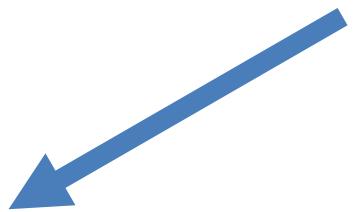
| Feature | SX530                               | RX100                    |
|---------|-------------------------------------|--------------------------|
| Zoom    | 50x                                 | 3.6x                     |
| Prize   | \$249.99                            | \$399.99                 |
| Weight  | 15.68 ounces                        | 7.5 ounces               |
| Prefer  | <input checked="" type="checkbox"/> | <input type="checkbox"/> |



| Feature    | TG-4                     | G9                                  |
|------------|--------------------------|-------------------------------------|
| Waterproof | Yes                      | No                                  |
| Prize      | \$249.99                 | \$399.99                            |
| Weight     | 7.36 lb                  | 7.5 lb                              |
| Prefer     | <input type="checkbox"/> | <input checked="" type="checkbox"/> |



| Feature    | TG-4                                | Galaxy 2                 |
|------------|-------------------------------------|--------------------------|
| Waterproof | Yes                                 | No                       |
| Prize      | \$249.99                            | \$399.99                 |
| Viewfinder | Electronic                          | Optical                  |
| Prefer     | <input checked="" type="checkbox"/> | <input type="checkbox"/> |



Estimate of  
preference  
parameter

- Today: Minimize **variance** of **parameter** estimates

# Parametric Model = Logistic Regression



| Feature | Chewbacca                           | BB-8                     |
|---------|-------------------------------------|--------------------------|
| Wookiee | Yes                                 | No                       |
| Droid   | No                                  | Yes                      |
| Blaster | Yes                                 | No                       |
| Prefer? | <input checked="" type="checkbox"/> | <input type="checkbox"/> |

$$x^1 \quad x^2$$

Product profile

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = x^2$$

MNL Random Linear Utility

$$U_j = \underbrace{\beta \cdot x^j}_{\sum_{j=1}^d \beta_i x_i^j} + \epsilon_j$$

$$z = x^1 - x^2$$

Question:

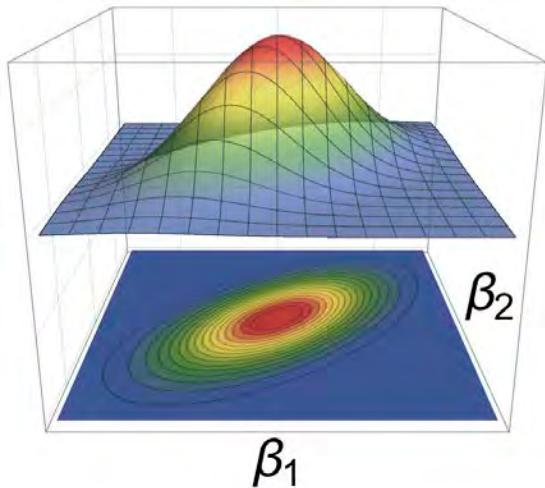
$$x^1 \succ x^2 \Leftrightarrow U_1 \text{ “>” } U_2$$

$$\Leftrightarrow \beta \cdot z \text{ “>” } 0$$

$$\mathbb{P}(x^1 \succ x^2 | \beta) = \frac{1}{1 + e^{-\beta \cdot z}}$$

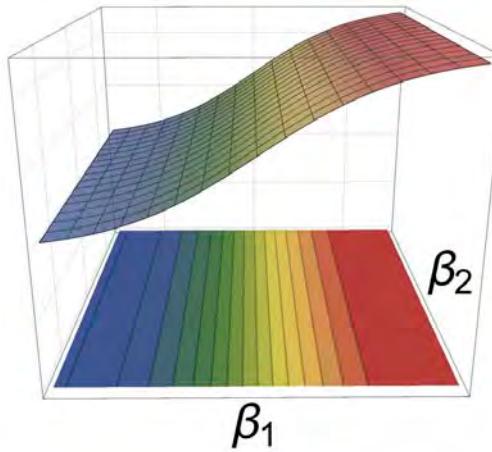
# Bayesian Model with Normal Prior

Prior distribution



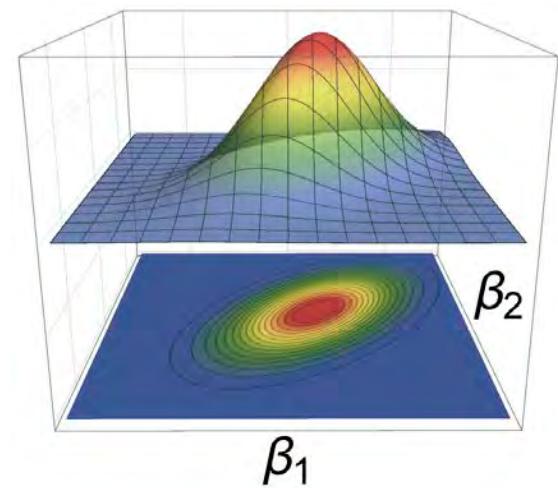
$$\beta \sim N(\mu, \Sigma)$$

Answer likelihood



$$L(y | \beta, z)$$

Posterior distribution



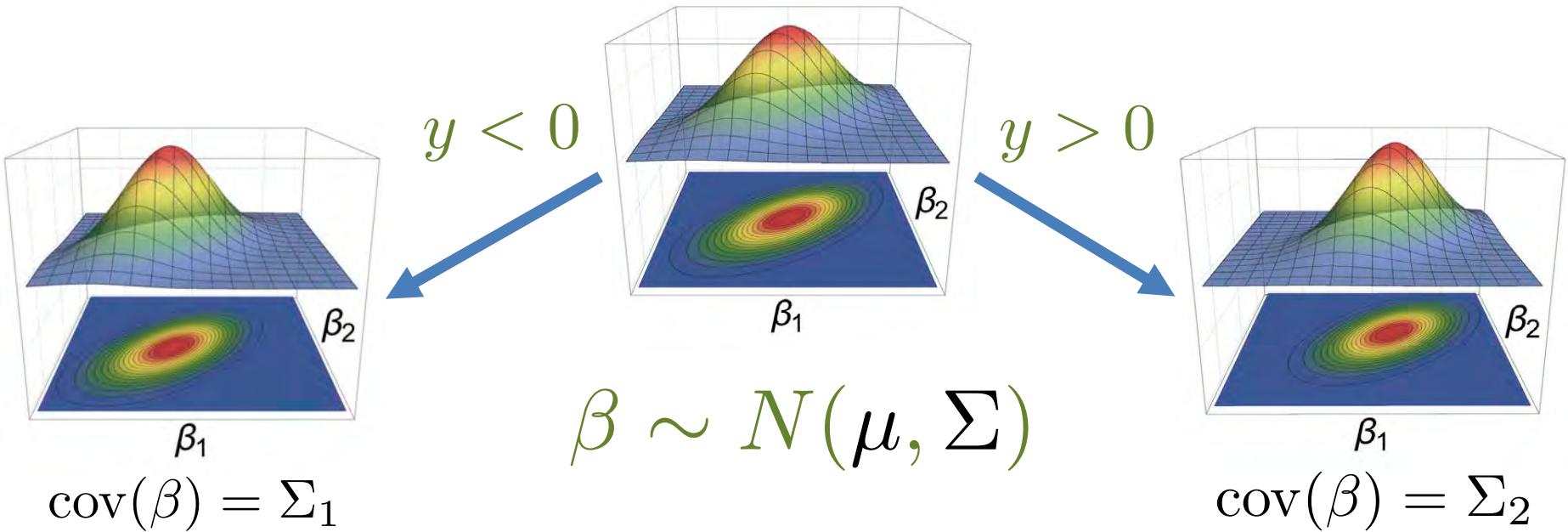
$$g(\beta | y, z)$$

$$y = \text{sign}(\beta \cdot z) \quad L(y | \beta, z) = (1 + e^{-y\beta \cdot z})^{-1}$$

$$g(\beta | y, z) \propto \phi(\beta ; \mu, \Sigma) L(y | \beta, z)$$

# D-Error and Expected Posterior Variance

$$f(\mathbf{z}, \mu, \Sigma) := \mathbb{E}_{y, \beta} \left\{ (\det \text{cov}(\beta | y, \mathbf{z}))^{1/m} \right\}$$



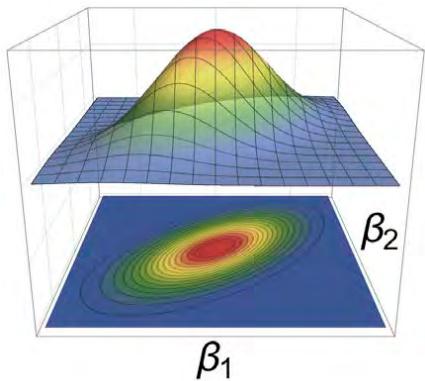
$$\min_{\mathbf{z} \in \{-1, 0, 1\}^n \setminus \{\mathbf{0}\}} f(\mathbf{z}, \mu, \Sigma)$$

- $f(\mathbf{z}, \mu, \Sigma)$  is hard to evaluate, non-convex and  $n$  large

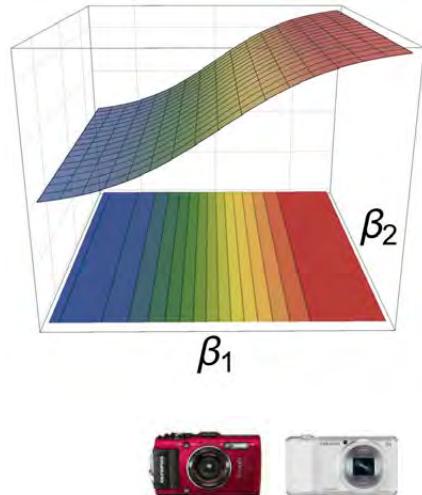
# 1<sup>st</sup> Step: Moment-Matching Approximate Bayes

Answer likelihood

Prior distribution

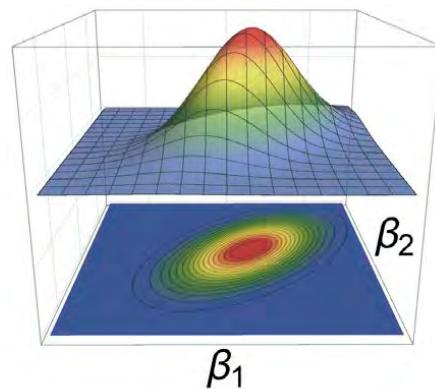


$$\beta \sim N(\mu^i, \Sigma^i)$$



| Feature    | TG-4         | Galaxy 2 |
|------------|--------------|----------|
| Waterproof | Yes          | No       |
| Prize      | \$249.99     | \$399.99 |
| Viewfinder | Electronical | Optical  |
| Prefer     | ✓            | □        |

Posterior distribution



$$\beta \stackrel{approx.}{\sim} N(\mu^{i+1}, \Sigma^{i+1})$$

- $\mu^{i+1} = \mathbb{E}(\beta | y, x^1, x^2)$
- $\Sigma^{i+1} = \text{cov}(\beta | y, x^1, x^2)$

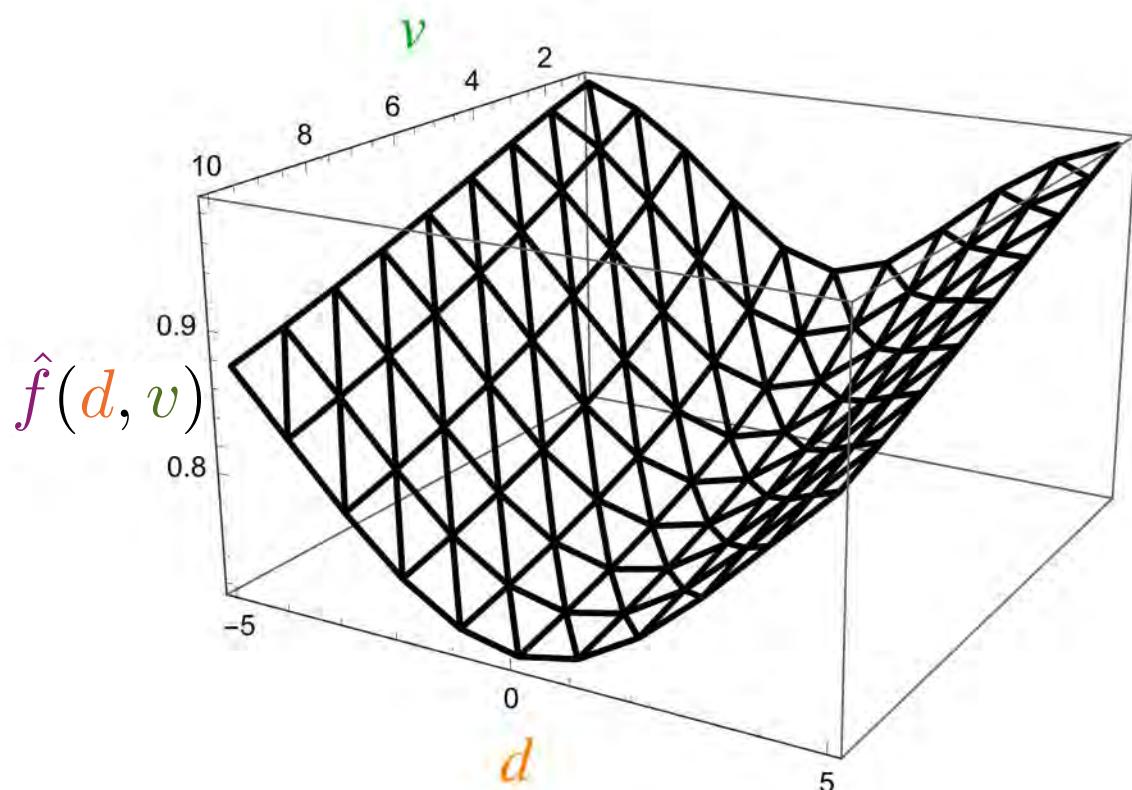
- Linear Algebra + 1-d numerical integration (e.g. BDA3)

## 2<sup>nd</sup> Step: More Linear Algebra from V. and S. '16

- D-efficiency  $f(z)$  = Non-convex function  $f(d, v)$  of

mean:  $d := \mu \cdot z$

variance:  $v := z' \cdot \sum \cdot z$



Can evaluate  $f(d, v)$  with 1-dim integral 😊

Piecewise Linear (PWL)  
Interpolation  $\hat{f}(d, v)$

Balances known criteria:

- minimize mean of question (no clear expected answer)
- maximize variance of question (uncertainty in expected answer)

### 3<sup>rd</sup> Step: “Almost” Direct Linear MIP Formulation

$$z = x^1 - x^2$$

MIP formulation for PWL function

$\min$

$$\hat{f}(d, v)$$



$s.t.$

$$\mu \cdot (x^1 - x^2) = d$$

$$(x^1 - x^2)' \cdot \sum \cdot (x^1 - x^2) = v$$

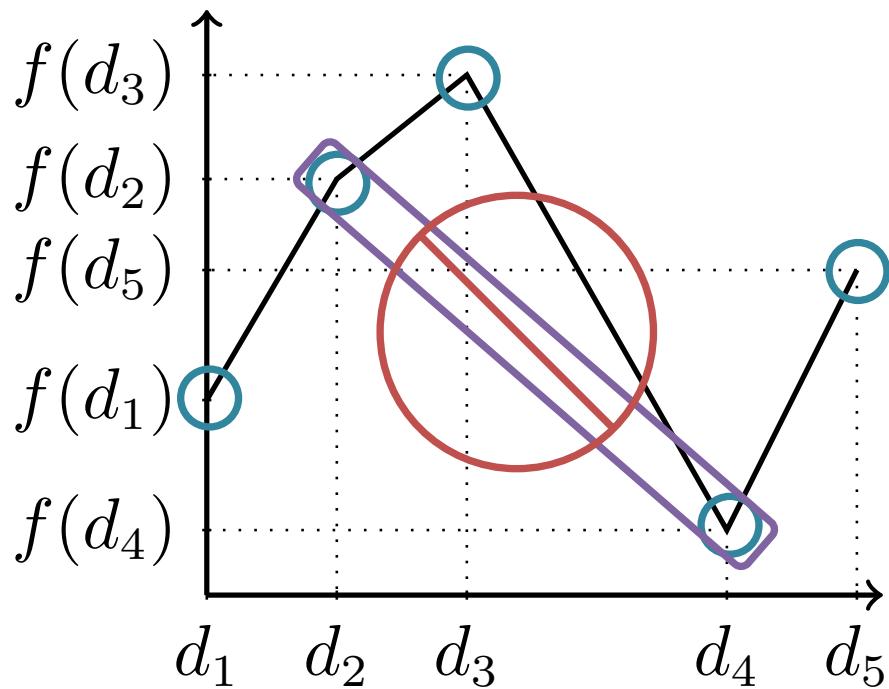


$$\text{linearize } x_i^k \cdot x_j^l \quad \|x^1 - x^2\|_2^2 \geq 1 \quad (x^1 \neq x^2)$$

$$x^1, x^2 \in \{0, 1\}^n$$

# Simple Formulation for Univariate Functions

$$z = f(x)$$



Size =  $O(\# \text{ of segments})$

Non-Ideal: Fractional Extreme Points

$$\begin{pmatrix} x \\ z \end{pmatrix} = \sum_{j=1}^5 \begin{pmatrix} d_j \\ f(d_j) \end{pmatrix} \lambda_j$$
$$1 = \sum_{j=1}^5 \lambda_j, \quad \lambda_j \geq 0$$

$$y \in \{0, 1\}^4, \quad \sum_{i=1}^4 y_i = 1$$

$$0 \leq \lambda_1 \leq y_1$$

$$0 \leq \lambda_2 \leq y_1 + y_2$$

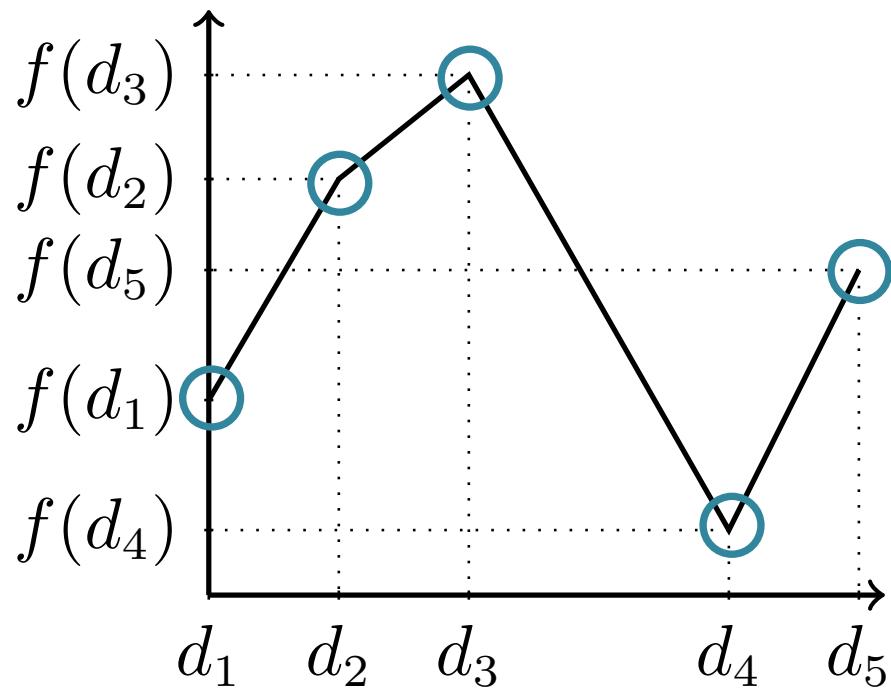
$$0 \leq \lambda_3 \leq y_2 + y_3$$

$$0 \leq \lambda_4 \leq y_3 + y_4$$

$$0 \leq \lambda_5 \leq y_4$$

# Advanced Formulation for Univariate Functions

$$z = f(x)$$



Size =  $O(\log_2 \# \text{ of segments})$

Ideal: Integral Extreme Points

Significant computational advantage

$$\begin{aligned} \binom{x}{z} &= \sum_{j=1}^5 \binom{d_j}{f(d_j)} \lambda_j \\ 1 &= \sum_{j=1}^5 \lambda_j, \quad \lambda_j \geq 0 \end{aligned}$$

$$y \in \{0, 1\}^2$$

$$0 \leq \lambda_1 + \lambda_5 \leq 1 - y_1$$

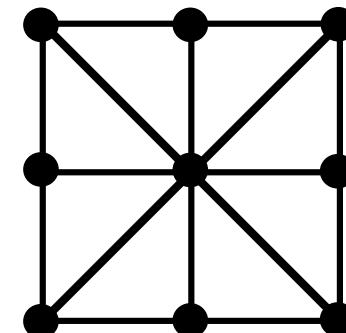
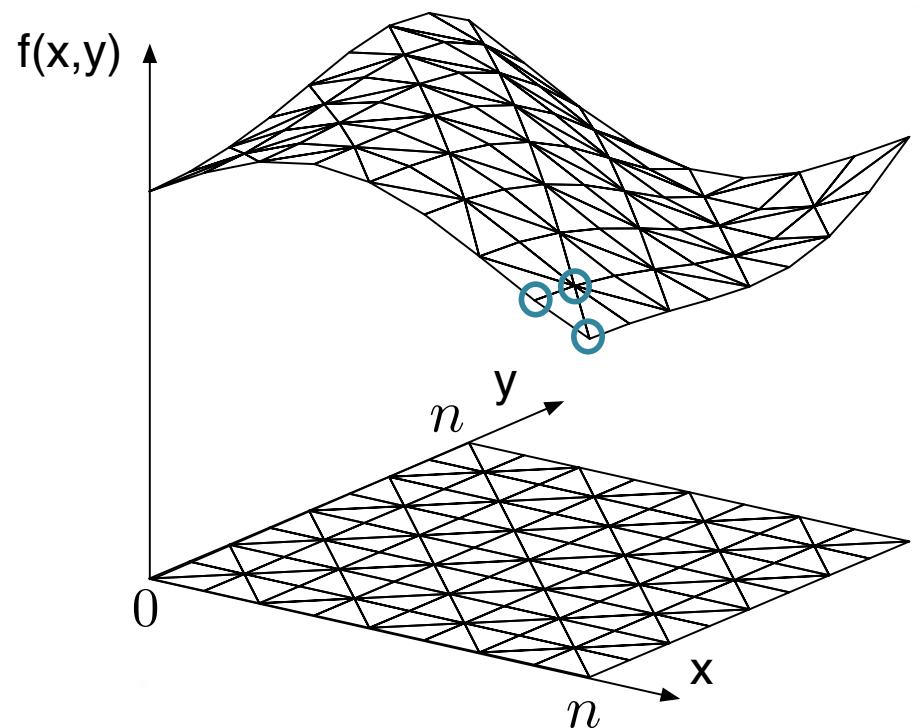
$$0 \leq \lambda_3 \leq y_1$$

$$0 \leq \lambda_4 + \lambda_5 \leq 1 - y_2$$

$$0 \leq \lambda_1 + \lambda_2 \leq y_2$$

# Technique Also Works for Multivariate Functions

- Union Jack triangulation  
(V. and Nemhauser, 2011)
  - Size =  $4 \lceil \log_2 n \rceil + 2$
- For general triangulations  
(Huchette and V., 2016, 2017)
  - Size  $\leq 4 \lceil \log_2 n \rceil + 6$
  - Based on finding a bi-clique cover of an auxiliary graph
    - Can use a MIP to find the smallest formulation!



# Easy to Build through **julia** & **JuMP**

- PiecewiseLinearOpt.jl (**Huchette** and V. 2017)

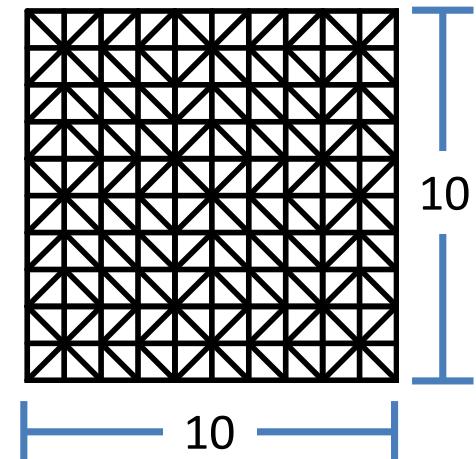
$$\min \quad \exp(x + y)$$

s.t.

$$x, y \in [0, 1]$$

Automatically select  $\Delta$

Automatically construct  
formulation (easily chosen)



```
using JuMP, PiecewiseLinearOpt
m = Model()
@variable(m, x)
@variable(m, y)

z = piecewiselinear(m, x, y, 0:0.1:1, 0:0.1:1, (u,v) -> exp(u+v))
@objective(m, Min, z)
```

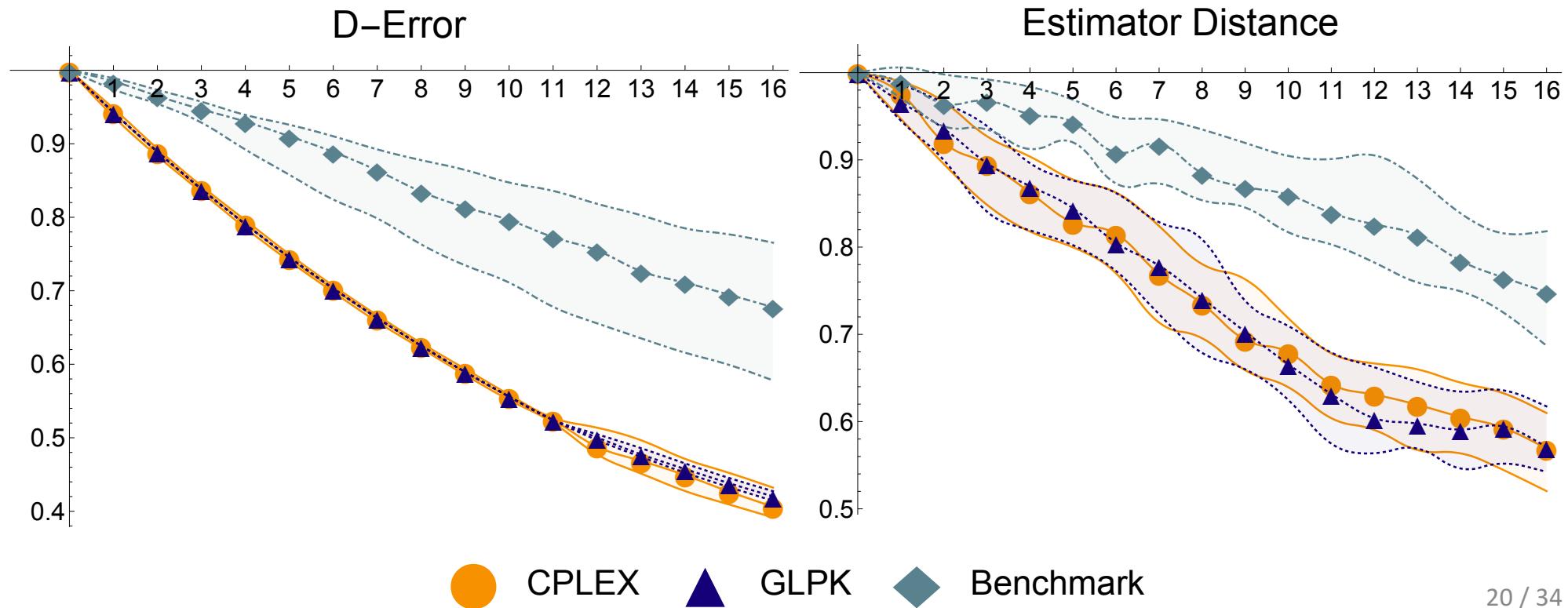
# Easy to Build through **julia** & **JuMP**

---

```
function getquestion(mu,S,variancefuction)
    n = size(S,1)
    m = Model()
    # define variables for linearization
    @variable(m, 0 <= x[1:n] <= 1, Int)
    @variable(m, 0 <= y[1:n] <= 1, Int)
    # x ≠ y
    @constraint(m, linquad(m,(x-y)·(x-y)) >= 1)
    # v = x-y, β ~ N(mu,S), v·β ~ N(mu_v,σ²), μ_v = mu·v, σ² = v'·S·v
    @variable(m, mu_v)
    @constraint(m, mu_v == mu·(x-y) )
    @variable(m, σ² >=0)
    @constraint(m, σ² == linquad(m,(x-y)·(S*(x-y))))
    # (x-y)'·S·(x-y) <= eigmax(S) |||x-y|||₂ <= eigmax(S)*n
    σ² = eigmax(S)*n
    # (x-y)'·S·(x-y) >= eigmin(S) |||x-y|||₂ >= eigmin(S) ( x ≠ y )
    σ² = eigmin(S)
    μ_v = norm(mu,1)
    μ_vnpoints = 2^k - 1
    μ_vpoints = 0:μ_v/μ_vnpoints:μ_v+(μ_v/μ_vnpoints)/2
    σ²range = σ² - σ²
    σ²npoints = 2^k-1
    σ²points = σ²:σ²range/σ²npoints:σ²+(σ²range/σ²npoints)/2
    pw1 = BivariatePWLFunction(μ_vpoints, σ²points, (μ_v,σ²) -> variancefuction(μ_v,sqrt(σ²)))
    obj = piecewiselinear(m, μ_v, σ², pw1)
    @objective(m, Min, obj )
    status = solve(m)
    return [ round(Int64,getvalue(x)), round(Int64,getvalue(y))]
end
```

# MIP v/s Best Benchmark (Toubia et al. '03,'04)

- 16 questions, 2 options, 12 features, 100 individual  $\beta^*$  sampled from known prior  $N(\mu, \Sigma)$
- Best **benchmark** v/s MIP + Moment Matching
- **CPLEX:  $\leq 1$  s (0.2 s Avg.), GLPK:  $\leq 5$  s ( 1.7 s Avg.)**



# Easy To Add Questionnaire Rules



Product profile

$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = x^2$

| Feature | Chewbacca                           | BB-8                     |
|---------|-------------------------------------|--------------------------|
| Wookiee | Yes                                 | No                       |
| Droid   | No                                  | Yes                      |
| Blaster | Yes                                 | No                       |
| Prefer? | <input checked="" type="checkbox"/> | <input type="checkbox"/> |

- Realism is important: Wookiees are not Droids!

$$x_{\text{Wookie}}^1 + x_{\text{Droid}}^1 \leq 1$$

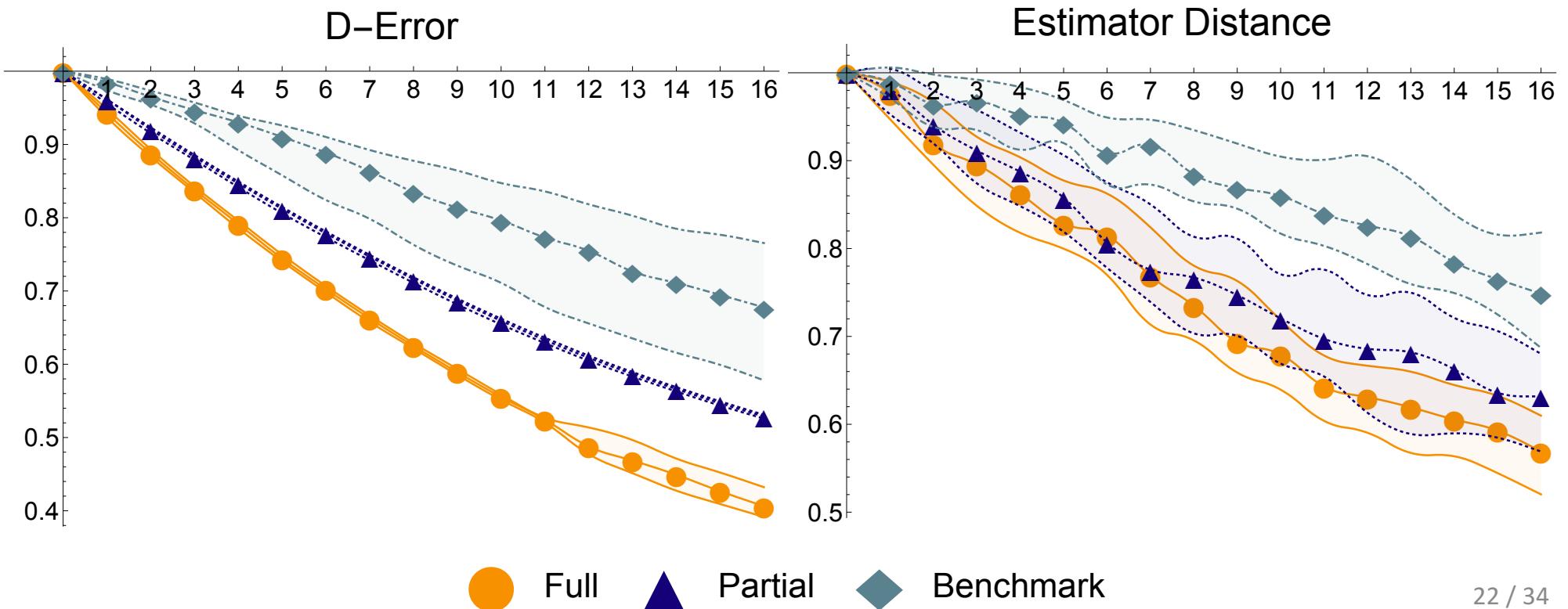
- Partial Profiles:

- Limit # of feature differences and assume those not shown are the same (e.g. both are members of the resistance)

$$\|x^1 - x^2\|_1 \leq 3$$

# Full v/s Partial Profiles ( # Feature Differences )

- 16 questions, 2 options, 12 features, 100 individual  $\beta^*$  sampled from known prior  $N(\mu, \Sigma)$
- Best **benchmark** v/s MIP + Moment Matching (CPLEX)
- Full:  $\leq 1$  s (0.2 s Avg.), Partial (5 diff.):  $\leq 66$  s ( 8 s Avg.)



# MIP and Causal Inference

Joint work with Magdalena Bennett and  
Jose Zubizarreta

# Educational Impact of 2010 Chilean Earthquake

---

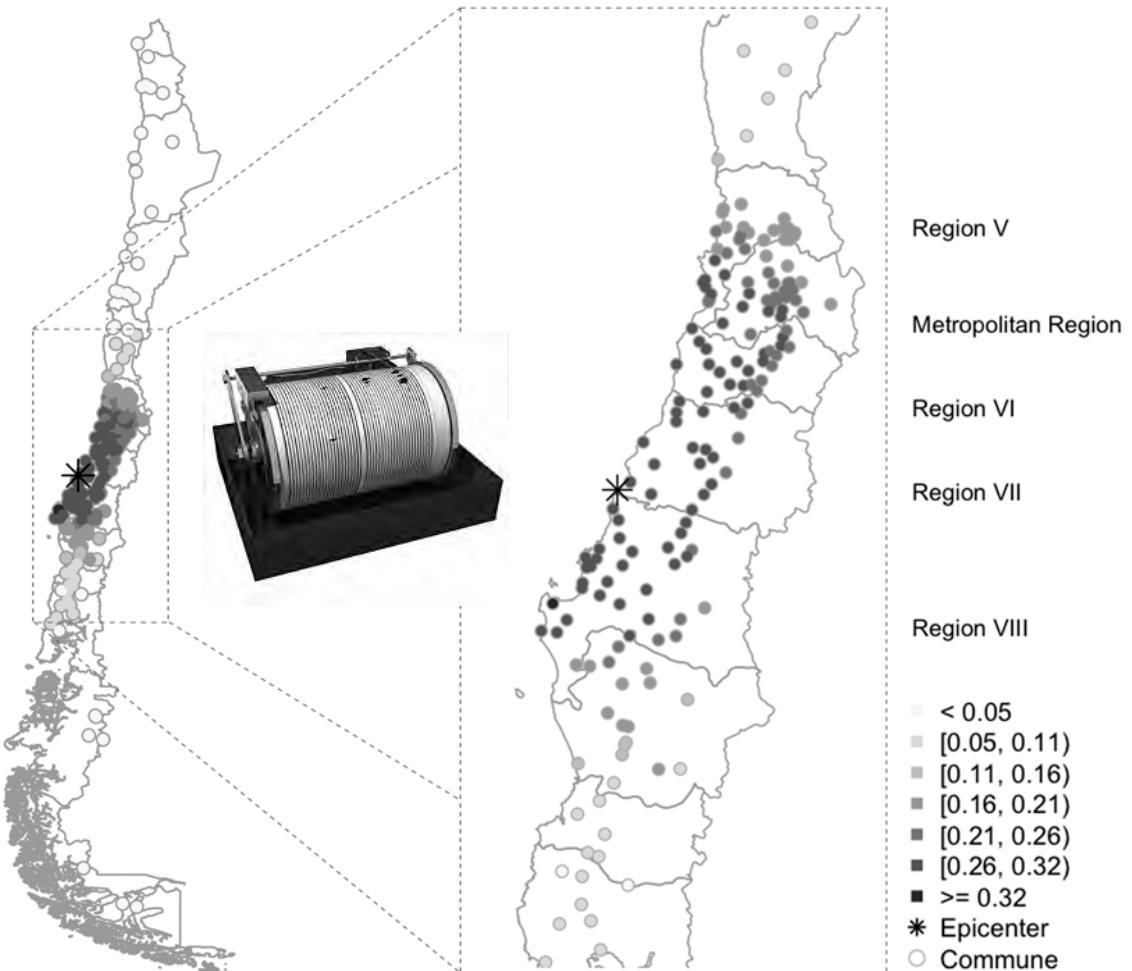


- 6th Strongest in Recorded History (8.8)



- Impact on Educational Achievement (PSU Scores)?

# Very High Quality Data is Available



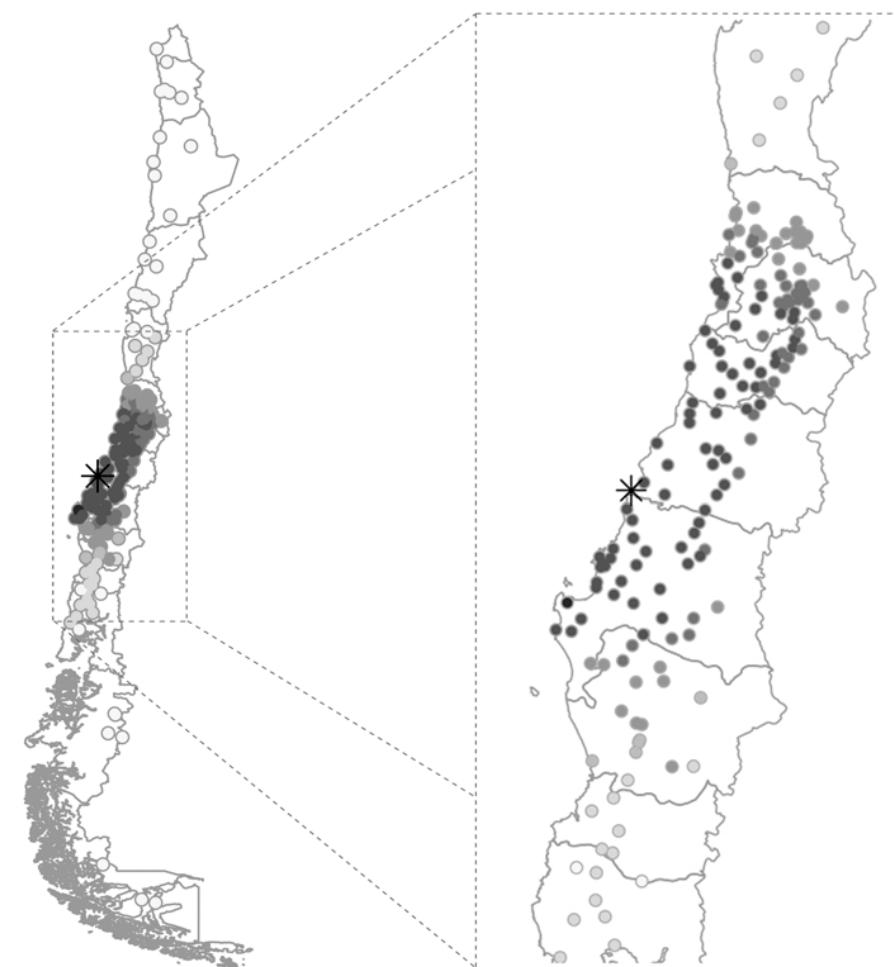
## Earthquake Intensity

## Test Scores and Demographic Info

# Covariate Balance Important for Inference

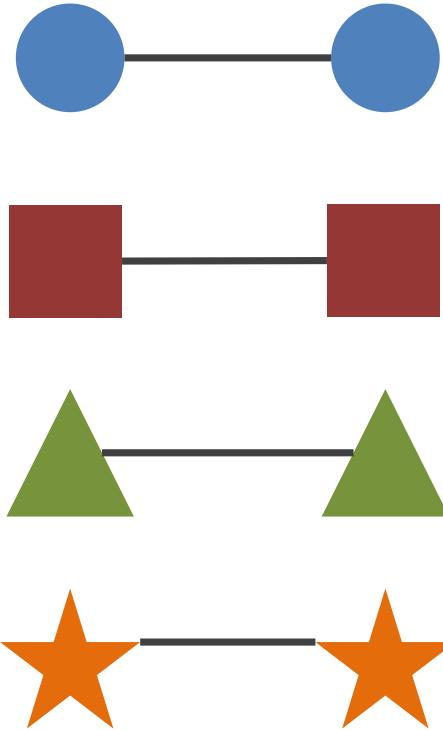
- Dose 1 = Control = Affected by Earthquake
- Dose 2 = Treatment = Not affected by Earthquake

| Covariate          | Dose |     |
|--------------------|------|-----|
|                    | 1    | 2   |
| Gender             |      |     |
| Male               | 462  | 462 |
| Female             | 538  | 538 |
| School SES         |      |     |
| Low                | 75   | 75  |
| Mid-low            | 327  | 327 |
| Medium             | 294  | 294 |
| Mid-high           | 189  | 189 |
| High               | 115  | 115 |
| Mother's education |      |     |
| Primary            | 335  | 335 |
| Secondary          | 426  | 426 |
| Technical          | 114  | 114 |
| College            | 114  | 114 |
| Missing            | 11   | 11  |
| ⋮                  |      |     |



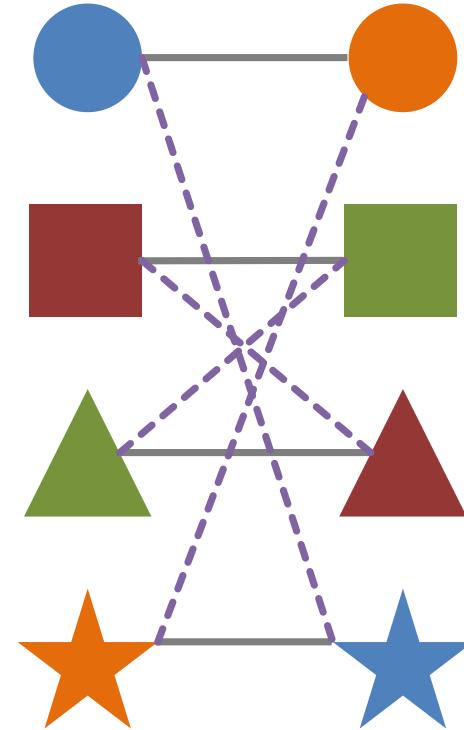
# Traditional Matching: Exact v/s Fine Balance

---



## Exact Matching

Match units with same category in all covariates



## Fine Balance

Different matches for different covariates

# Matching v/s MIP

---

- Matching: # Variables = (# treated ) × (# controls)
- Simple MIP:
  - Just count & balance units in each category/covariate
  - # Variables = (# treated ) + (# controls)
  - Can use known results (Balas and Pulleyblank , 1983) to show MIP formulation is as strong as matching formulation
  - Can show MIP formulation is integral for 2 covariates
  - Problem is NP-hard for > 2 covariates
  - Usually very fast solve times: cuppa coffee time  $\approx$  5 min
  - More flexible.... doses and representability!!

# Multiple Doses + Representation

---

- Base, Medium and High

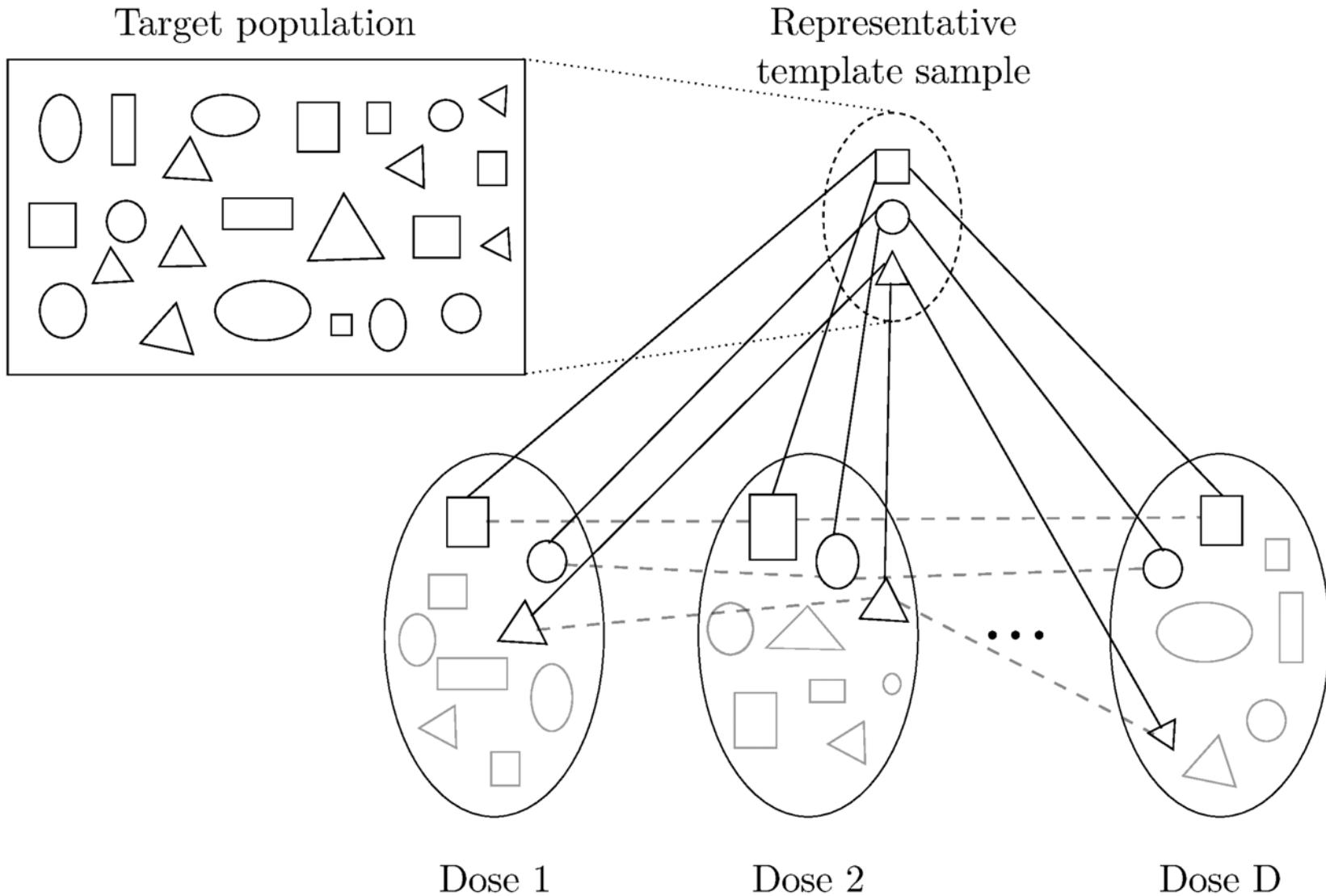
| Covariate          | Dose |     |     |
|--------------------|------|-----|-----|
|                    | 1    | 2   | 3   |
| Gender             |      |     |     |
| Male               | 462  | 462 | 462 |
| Female             | 538  | 538 | 538 |
| School SES         |      |     |     |
| Low                | 75   | 75  | 75  |
| Mid-low            | 327  | 327 | 327 |
| Medium             | 294  | 294 | 294 |
| Mid-high           | 189  | 189 | 189 |
| High               | 115  | 115 | 115 |
| Mother's education |      |     |     |
| Primary            | 335  | 335 | 335 |
| Secondary          | 426  | 426 | 426 |
| Technical          | 114  | 114 | 114 |
| College            | 114  | 114 | 114 |
| Missing            | 11   | 11  | 11  |
| :                  |      |     |     |

- Whole population

| Covariate          | Population | Template |
|--------------------|------------|----------|
| Father's education |            |          |
| Secondary          | 0.39       | 0.40     |
| Technical          | 0.09       | 0.09     |
| College            | 0.15       | 0.14     |
| Missing            | 0.05       | 0.04     |
| Mother's education |            |          |
| Secondary          | 0.41       | 0.43     |
| Technical          | 0.13       | 0.12     |
| College            | 0.12       | 0.12     |
| Missing            | 0.01       | 0.01     |
| Household income   |            |          |
| 100-200            | 0.26       | 0.26     |
| 200-400            | 0.30       | 0.31     |
| 400-600            | 0.13       | 0.12     |
| 600-1400           | 0.13       | 0.14     |
| 1400 or more       | 0.09       | 0.09     |
| Missing            | 0.01       | 0.02     |

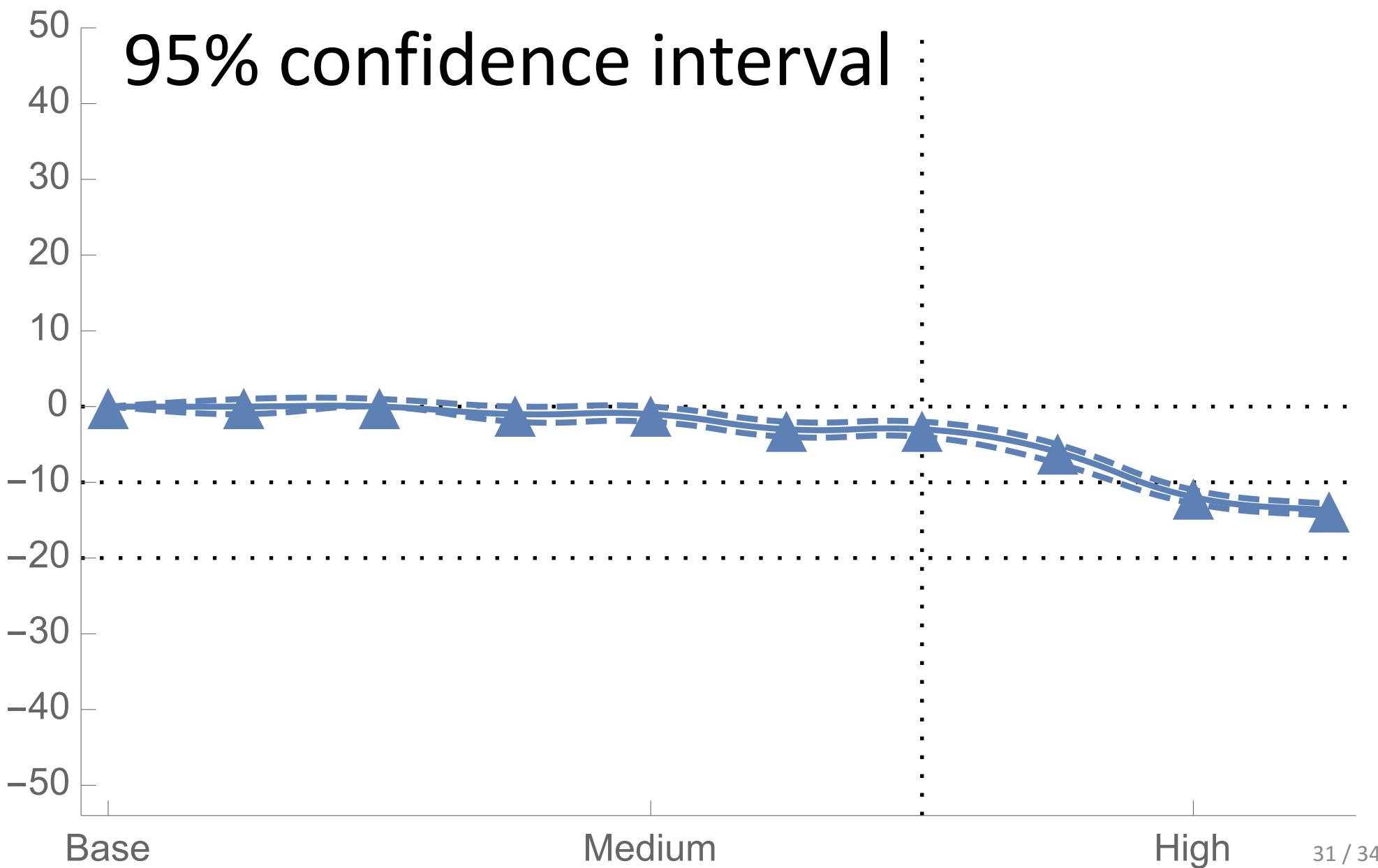
# Template + Multiple Doses

---

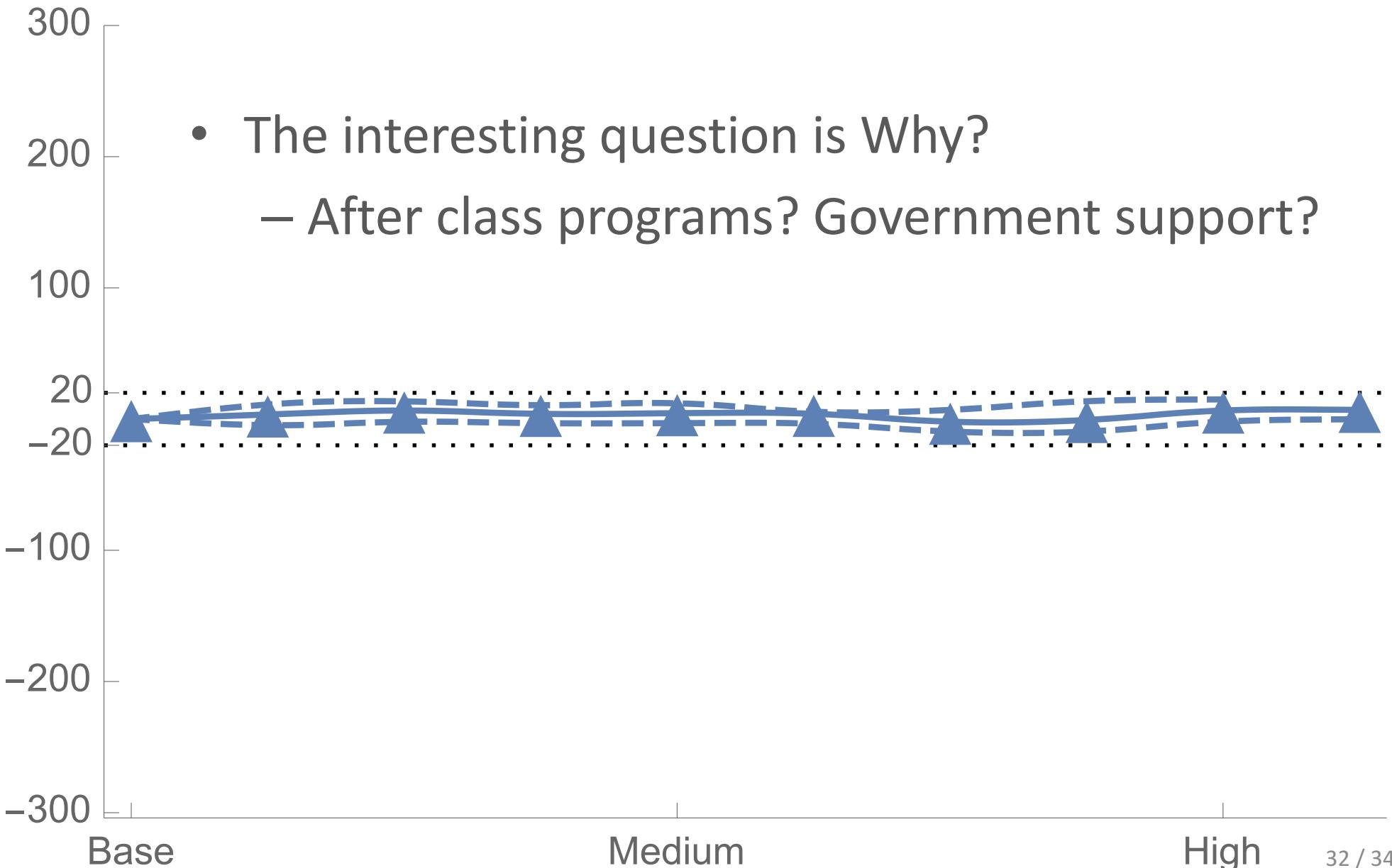


# Relative (To no Quake) Attendance Impact [%]

# 95% confidence interval

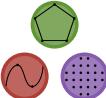


## Relative (To no Quake) PSU Score Impact (150–850)



# Summary

---

- Advances in MIP
  - Advanced Formulations
  - Advanced Solvers
  - Easy Access Through  JuMP
- Direct advantage (Choice Based Conjoint Analysis)
  - Real-time and versatile adaptive questionnaires
  - Cut number of questions in half
  - 20% improvement in estimated parameter quality
  - Market-share predictions cut in half
- Indirect advantage (Causal Inference)
  - Good and flexible formulations can bring you back to solving the problem you really wanted to solve