Multiple Hypothesis Testing in Conjoint Analysis

Yuki Shiraito

University of Michigan

Formal and Quantitative Seminar Department of Political Science The University of North Carolina at Chapel Hill

March 11, 2022

Joint work with Guoer Liu (U-M)

Conjoint Analysis

Conjoint Design

	Immigrant 1	Immigrant 2						
Prior Trips to the U.S.	Entered the U.S. once before on a tourist visa	Entered the U.S. once before on a tourist visa						
Reason for Application	Reunite with family members already in U.S.	Reunite with family members already in U.S.						
Country of Origin	Mexico	Iraq						
Language Skills	During admission interview, this applicant spoke fluent English	During admission interview, this applicant spoke fluent English						
Profession	Child care provider	Teacher						
Job Experience	One to two years of job training and experience	Three to five years of job training and experience						
Employment Plans	Does not have a contract with a U.S. employer but has done job interviews	Will look for work after arriving in the U.S.						
Education Level	Equivalent to completing two years of college in the U.S.	Equivalent to completing a college degree in the U.S.						
Gender	Female	Male						

Please read the descriptions of the potential immigrants carefully. Then, please indicate which of the two immigrants you would personally prefer to see admitted to the United States.

AMCE: test multiple causal hypotheses simultaneously

Classic Conjoint Results

Gender: female male	
Education: no formal 4th grade 8th grade high school two-year college college degree graduate degree	•==- ====
Language: fluent English broken English tried English but unable used interpreter	+ → •
Origin: Germany France Maxico Philippines Poland Iradia Ohina Sudan Sudan Somalia Iraq	
Profession: janitor waiter child care provider gardener financial analyst construction worker teacher computer programmer nurse research scientist doctor	
Job experience: none 1-2 years 3-5 years 5+ years	·
Job plans: contract with employer interviews with employer will look for work no plans to look for work	+ + +
Application reason: reunite with family seek better job escape persecution	
Prior trips to U.S.: never once as tourist many times as tourist six months with family once w/o authorization	
	Change in Pr(Immigrant Preferred for Admission to U.S.)

Hainmueller et. al. (2014), p.21

Liu and Shiraito (U-M)

Multiple Testing in Conjoint Analysis

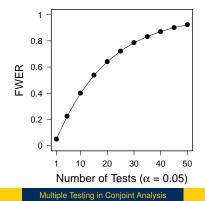
Multiple Hypothesis Testing

- Test one hypothesis, $\alpha \equiv \mathbb{P}(\text{Reject null} \mid \text{Null is true}) = 0.05$
- Test ten hypotheses simultaneously with $\alpha = 0.05$

 $\mathsf{FWER} \equiv \mathbb{P}(\mathsf{At \ least \ one \ null \ is \ rejected} \mid \mathsf{All \ nulls \ are \ true})$

$$=$$
1 - (1 - α)¹⁰ \approx .4

• Family-Wise Error Rate as the Number of Tests Increases



Liu and Shiraito (U-M)

Number of Hypotheses in Conjoint Analysis: 41

Gender:		
female	•	
male		
Education:		
no formal		
4th grade		
401 grade		
8th grade		
high school		
two-year college		
college degree		
graduate degree	·	
° °		
Language:		
fluent English		
holice English		
broken English tried English but unable		
used interpreter		
used interpreter	·	
Origin:		
Origin: Germany		
Mexico	· · · · · · · · · · · · · · · · · · ·	
Philippines	· · · · · · · · · · · · · · · · · · ·	
Poland		
India		
China		
Sudan		
Sudan		
Somalia	· · · · · · · · · · · · · · · · · · ·	
Iraq	·	
Profession:		
janitor	• •	
waiter		
child care provider		
critic care provider		
gardener financial analyst		
gardener financial analyst		
gardener financial analyst construction worker		
gardener financial analyst construction worker teacher		
gardener financial analyst construction worker teacher computer programmer		
gardener financial analyst construction worker teacher computer programmer nurse		
gardener financial analyst construction worker teacher computer programmer nurse research scientist		
gardener financial analyst construction worker teacher computer programmer nurse		
gardener financial analyst construction worker teacher computer programmer nurse research scientist doctor		
gardener financial analyst construction worker teacher computer programmer nurse research scientist doctor		
gardener ' financial analyst construction worker teacher computer programmer nurse research scientist doctor Job experience: none		
gardener ' financial analyst construction worker teacher computer programmer nurse research scientist doctor Job experience: none		
gardener ' financial analyst construction worker teacher computer programmer nurse research scientist doctor Job experience: none		<u>. </u>
gardener i financial analyst construction worker teacher computer programmer nurse research scientist doctor Job experience: none 1–2 years		
gardener ' financial analyst construction worker teacher computer programmer nurse research scientist doctor Job experience: none		
gardener i financial analyst construction worker teacher omputer programmer nurse research scientist doctor Job experience: none none 3-5 years 5+ years		
gardener ' financial analyst construction worker teacher computer programmor more 1-2 years 3-5 years 3-5 years 3-5 years 3-5 years		
gardener ' financial analyst construction worker teompier programmer nurse research scientist doctor Job experience: 1-2 years 3-5 years 5- years Job plans: contract with employer		
gardener i financial analyst construction worker teacher omputer programmer nurse rection accientist doctor Job experience: none 1–2 years 3–5 years 3–5 years 5– years 5– years interviews with employer interviews with employer		
gardrener i Inancial analyst construction worker Beacher nurse programmer nurse programmer nurse programmer dotter Job experience: 3-3 years 3-5 years 5-4 years 5-4 years 3-5 years 3-5 years 3-6 years 3-6 years 3-6 years 3-7 years 3-8 years 3-8 years 3-8 years 3-9 y	•	<u>.</u>
gardener i financial analyst construction worker teacher omputer programmer nurse rection accientist doctor Job experience: none 1–2 years 3–5 years 3–5 years 5– years 5– years interviews with employer interviews with employer	•	
gardrener i Inancial analyst construction worker Beacher nurse programmer nurse programmer nurse programmer dotter Job experience: 3-3 years 3-5 years 5-4 years 5-4 years 3-5 years 3-5 years 3-6 years 3-6 years 3-6 years 3-7 years 3-8 years 3-8 years 3-8 years 3-9 y	•	
gardener i financial analyst teacher innoviter teacher innoviter teacher innoviter nurse computer programmer nurse computer programmer teasearch scientist doctor no 2 years 3 - 5 years 3 - 5 years 5 + years contract with employer will lock for work mo plans to lock for work	•	
gardnear i financial ranjat inancial ranjat inancial ranjat computer programmer nursa doctor Job spreience: nonon 1–2 years 3–5 years 3–5 years 5–4 years 3–5 years 5–4 years 3–5 years 5–4 years 3–5 years 5–6 years 5–7 years 5–	•	
gardnear innancial rankyst construction worker computer programmer nurse computer programmer nurse computer programmer nurse doctor	•	
andrener i financial analyst teacher in worker teacher in worker computer programmer nurse computer programmer nurse doctor 1-2 years 3-5 years 3-5 years 3-5 years 3-5 years 3-5 years 3-5 years 4-1 guans to look for work Application renson, Application renson, seek better (ob)	•	<u>.</u>
gardnear innancial rankyst construction worker computer programmer nurse computer programmer nurse computer programmer nurse doctor	•	<u> </u>
gardener i financial aralyst financial aralyst teacher computer programmer research scientist doctor Job xoprience: none 1 – 5 years 5 + years 5 + years Job plans: contract with employer interviews with employer	•	
gardener ' gardener' vyst boarouitst vorker teacher compose programmer compose programmer doctor Job soprence: 10-2 years 3-5 years 5-5 years 5-5 years 5-5 years 5-5 years 5-5 years 10-0 kor work will book for work Application reason: no plans to look for work Application reason: no plans to look for work Application reason: no plans to look for work Application reason: Not for to to U.S.:	•	
gardnear ' gardnear ' constructor worker teacher onstructor programmer nurse research scientist doctor 1-2 years 3-5 years doctor 1-2 years 3-5 years doctor 1-2 years 3-6 years doctor 1-2 years 1-2 years 3-6 years doctor 1-2 years 1-2 years	•	<u> </u>
gardnear ' gardnear ' boostructor worker boostructor worker boostructor worker boostructor worker boostructor boost constructor work boostructor boost contract with employer contract with employer contract with employer boostructor boost for work will book for work mo plans to look for work boost for work mo plans to look for work boost boots for work boots for	•	
partern right of the second se	•	<u> </u>
particlem ¹ , particlem ¹ , constructors version constructors version constructors version constructors version constructors version constructors doctor doct	•	<u> </u>
partern right of the second se	•	
particlem ¹ , particlem ¹ , constructors version constructors version constructors version constructors version constructors version constructors doctor doct	•	2

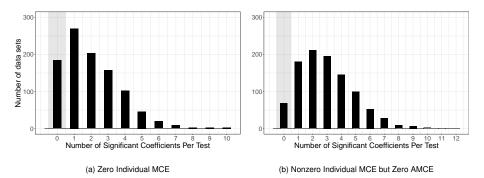
Change in Pr(Immigrant Preferred for Admission to U.S.)

Liu and Shiraito (U-M)

Multiple Testing in Conjoint Analysis

Quantifying the Problem by Simulations

- If AMCE is zero, in how many samples do you get false findings?
- Two scenarios for 41 attribute levels:
 - No individual effect
 - Nonzero individual effect, but zero average effect
- Number of samples for each number of false findings:



Correction Methods Overview

- Objective: contain false positive conclusions
- Trade-off: risk false negative conclusions
- Correction methods
 - Control family-wise error rate (FWER)
 - Bonferroni Correction
 - Control false discovery rate (FDR)
 - Benjamini-Hochberg Procedure
 - Control false discovery rate (FDR) & Reduce RMSE
 - Adaptive Shrinkage
- Proposal:



Motivation

Simulation

Reanalysis

Bonferroni Correction

- \bullet Controls FWER to α
- Procedure: set $\alpha^* = \frac{\alpha}{\# \text{ of tests}}$ for each test
- Strength: easy to construct confidence intervals
- Shortcomings:

high risk of false negative conclusions ambiguous definition of "total number of tests"

Benjamini-Hochberg Procedure

Controls FDR:

$$\mathbb{E}\left[\frac{\text{\# of false discoveries}}{\text{\# of total discoveries}}\right] \le \alpha$$

- Solution:
 - Rank p-values from smallest to largest
 - Provide the null up to the largest p-value such that

$$p \leq \frac{\operatorname{rank} \operatorname{of} p}{\# \operatorname{of} \operatorname{tests}} \alpha$$

- Strength: less susceptive to false negative conclusion
- Shortcomings:
 - sensitive to pre-specified FDR no uncertainty measures

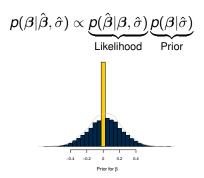
Liu and Shiraito (U-M)

Motivation

Simulation

Adaptive Shrinkage

• Regularizes β by placing a *spike-and-slab* prior



- Procedure: empirical Bayes post-estimation procedure
- Strength:

transparent, flexible, credible interval more precise point estimates

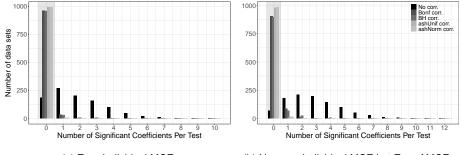
Liu and Shiraito (U-M)

Simulations

- Design matrix identical to Hainmueller et. al. (2014)
- Avoiding false positives: zero AMCE
 - No individual effect
 - 2 Nonzero individual effect, but zero average effect
- Avoiding both false positives and false negatives: nonzero AMCE
 - Only gender has effect (appendix)
 - 2 All levels of gender, education, English have effects

Reanalysis

Zero AMCE



(a) Zero Individual MCE

(b) Nonzero Individual MCE but Zero AMCE

Nonzero AMCE

					No.	of Fa	lse F	ositi	ves			
			0	1	2	3	4	5	6	7	8	9
	No corr.	9	2	8	3	1	4	1				
No. of True Positives	110 0011.	10	258	270	196	133	54	42	13	10	4	1
		8	38									
	Bonf corr.	9	305	6	2							
		10	623	25	1							
	DIL	8	4									
		9	47	25	4		1					
<u>No. of True Positives</u>		10	607	208	66	23	7	6	2			
		8	17	2								
	ashUnif corr.	9	160	26	4	1		1				
		10	620	127	30	6	5	1				
		8	21	2								
	ashNorm corr.	9	172	29	3	1	1					
		10	647	99	14	7	4					

Correct number of positives: 10

Reanalysis

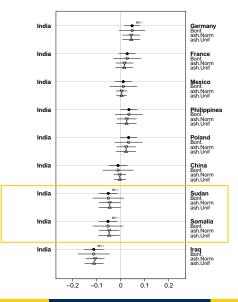
• Immigrants preferred by the U.S. public (Hainmueller et al. 2014)

- Focus on Country of Origin and Profession
- To show:



- How corrected results differ
- 2 ASh attains the middle
- Trading partners preferred in Vietnam (Spiker et al. 2016)
 - Focus on Military Ally and Environmental Standards
 - To show:
 - Bonf. and ASh recovers the null correctly
 - BH does not correct at all with few number of discoveries

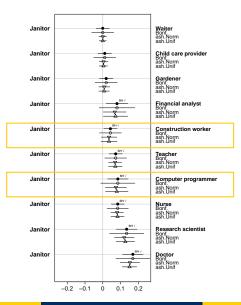
Country of Origin



Liu and Shiraito (U-M)

Multiple Testing in Conjoint Analysis

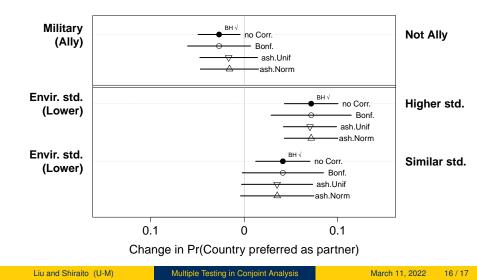
Profession



Liu and Shiraito (U-M)

Reanalysis

Selecting Trading Partners in Vietnam



Concluding Remarks

- Conjoint analysis inherently needs multiple hypothesis testing
- No correction ~> danger of false findings
- Correction methods
 - Bonferroni Correction (Most conservative)
 - Benjamini-Hochberg Procedure (Least conservative)
 - Adaptive shrinkage (middle-ground)



• Do correction, or you will get at least one false result

ASh Model

• Model: $\beta = (\beta_1, ..., \beta_J)$; est. $\hat{\beta}$, std.err $\hat{\sigma}$ $p(\beta|\hat{\beta}, \hat{\sigma}) \propto \underbrace{p(\hat{\beta}|\beta, \hat{\sigma})}_{\text{Likelihood}} \underbrace{p(\beta|\hat{\sigma})}_{\text{Prior}}$ $\beta_1, ..., \beta_J \stackrel{\textit{iid}}{\sim} g$

where

$$g(\cdot; \pi) = \pi_0 \delta_0(\cdot) + \sum_{k=1}^{K} \pi_k \mathcal{N}(\cdot; 0, \delta_k^2),$$

 $\sum_{k=0}^{K} \pi_k = 1 \quad \text{and} \quad \pi_k \ge 0$

Emprical Bayes estimates:

$$\hat{\pi} = \operatorname*{argmax}_{\pi} \prod_{j=1}^{J} \sum_{k=0}^{K} \pi_k \mathcal{N}(\hat{\beta}_j; \mathbf{0}, \delta_k^2 + \hat{s}_j^2)$$

Simulation Result: Only One Nonzero AMCE

		$\begin{array}{c ccccc} & \underline{\text{No. of False Positives}} \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \hline 1 & 230 & 290 & 215 & 123 & 69 & 42 & 19 & 9 & 3 \end{array}$												
			0	1	2	3	4	5	6	7	8			
	No corr.	1	230	290	215	123	69	42	19	9	3			
No. of True Positives	Bonf. corr.	1	966	32	2									
<u>INO. OF ITUE POSITIVES</u>	BH corr.	1	931	61	7	1								
	ashUnif corr.	1	996	4										
	ashNorm corr.	1	998	2										

 $\epsilon_i \stackrel{\textit{iid}}{\sim} \mathcal{N}(0, 0.01^2)$

Simulation Result: Only One Nonzero AMCE

						No	. of I	False	Posi	tive	s			10 11 1						
			0	1	2	3	4	5	6	7	8	9	10	11	12					
	No corr.	1	237	253	223	134	83	38	17	6	2	6			1					
No. of True Positives	Bonf. corr.	1	962	37	1															
INC. OF THE POSITIVES	BH corr.	1	930	55	7	5	1	1	1											
	ashUnif corr.	1	984	14	2															
	ashNorm corr.	1	987	12	1															

 $\epsilon_i \stackrel{\textit{iid}}{\sim} \mathcal{N}(0, 0.1^2)$

Simulation Result: Nonzero AMCE in Each Attribute

						of Fa									
			0	1	2	3	4	5	6	7	8	9	10	11	12
	No corr.	7		2				1							
	110 0011	8	10	22	27	16	22	8	2	3	1				
		9	118	194	179	169	86	58	39	19	13	7	2	1	1
		5	7	3											
No. of True Positives	Bonf corr.	6	77	5	2										
	Bolli corr.	7	244	15	7										
		8	396	37	5										
		9	180	20	2										
		6	5	2											
	BH corr.	7	37	15	5	1	1								
		8	147	89	36	11	4	1		3					
		9	321	187	75	35	12	8	1	3	1				
		6	12	3	1	1									
	ashUnif corr.	7	84	25	4	1	1								
		8	220	99	23	12	1	1							
		9	294	130	46	29	8	2	2		1				
		5		1											_
		6	11	5	2	1									
	ashNorm corr.	7	98	21	5	2									
		8	224	100	24	10	1	1							
		9	295	124	42	21	7	2	2	1					

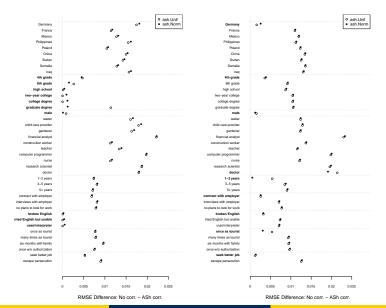
Figure: The true AMCE for each attribute has one significant levels I.

Simulation Result: Nonzero AMCE in Each Attribute

						 of J 										
			0	1	2	3	4	5	6	7	8	9	10	11	12	1
		6			5	7	4	4	1	1						
	No corr.	7			41	46	34	17	8	9	3					
		8			115	100	88	52	22	16	12	6	1		2	
		9			100	116	82	49	31	17	5	1	4			
		4		1	37											
		5		2	247	14	1									
	Bonf corr.	6		4	365	15	1									
		7		4	224	7	3	1								
No. of True Positives		8		2	63	2										
		9			7											
		4			3											
		5			32	4	2									
	BH corr.	6			106	28	7	4	2							
		7			212	70	17	8	1	1						
		8	_		229	82	38	9	7	2	1	1				
		9			77	34	13	5	3	2						
		4			2	1		1								
		5		1	52	13	4									
	ashUnif corr.	6		1	176	50	13	5								
		7			233	72	14	11	1	1						
		8			180	62	23	6	1	1	2					
		9			40	20	10	2	1	1						
		4			4			1								
		5		1	47	13	4									
	ashNorm corr.	6		1	174	49	11	3								
		7		234	71	17	8		1							
		8			187	63	23	7	1	2	1					
		9			43	20	11	1	2							

Figure: The true AMCE for each attribute has one significant levels II. The standard deviation for the reference category of *Job Experience* is four times larger.

Simulation Result: ASh RMSE



Liu and Shiraito (U-M)

Multiple Testing in Conjoint Analysis