

• Now, we observe a probability of choosing each alternative

Stated probabilistic choice models

• Now, we observe a probability of choosing each alternative

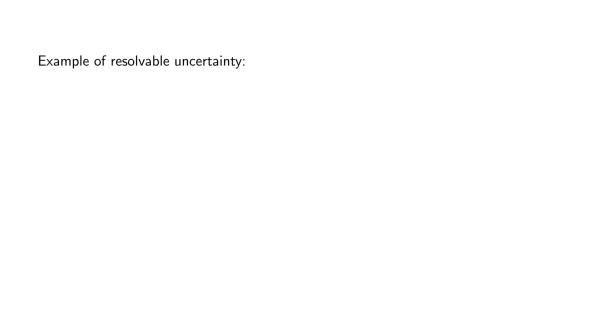
• What does this information represent? It represents the person's uncertainty

Stated probabilistic choice models

- Now, we observe a probability of choosing each alternative
- What does this information represent? It represents the person's uncertainty
- Specifically, it represents resolvable uncertainty:

Stated probabilistic choice models

- Now, we observe a probability of choosing each alternative
- What does this information represent? It represents the person's uncertainty
- Specifically, it represents resolvable uncertainty:
 - uncertainty about unspecified attributes or states of the world in which choices will ultimately be made



Exampl	e of	reso	lvable	e uncer	tainty
--------	------	------	--------	---------	--------

• Q. "Will you go for Mexican or Thai food on Friday night?"

Example of	of resolv	able unce	ertainty:
------------	-----------	-----------	-----------

• Q. "Will you go for Mexican or Thai food on Friday night?"

• A. "Well, I'm not sure what mood I'll be in that day, but probably Mexican"

Example of resolvable uncertainty:

- Q. "Will you go for Mexican or Thai food on Friday night?"
- A. "Well, I'm not sure what mood I'll be in that day, but probably Mexican"

• Q. "What do you mean by 'probably'?"

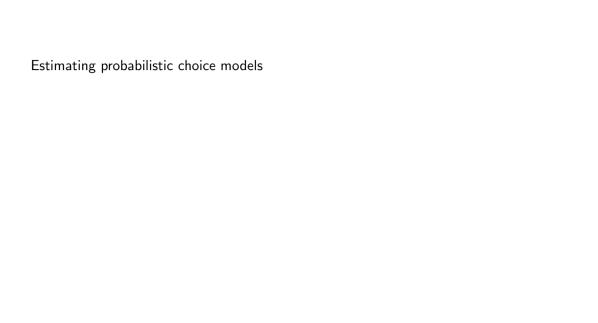
Example of resolvable uncertainty:

- Q. "Will you go for Mexican or Thai food on Friday night?"
- A. "Well, I'm not sure what mood I'll be in that day, but probably Mexican"
- Q. "What do you mean by 'probably'?"
- A. "A 78% chance"

Example of resolvable uncertainty:

- Q. "Will you go for Mexican or Thai food on Friday night?"
- A. "Well, I'm not sure what mood I'll be in that day, but probably Mexican"
- Q. "What do you mean by 'probably'?"
- A. "A 78% chance"

If the person thinks there is no such uncertainty, then they can report $\it p=0$ or $\it p=1$



Estimating probabilistic choice models
How do we proceed with estimation when y is itself a probability (not $0/1$)?

Estimating	probabilistic	choice	models

We invert the logit formula

How do we proceed with estimation when y is itself a probability (not 0/1)?

Estimating probabilistic choice models

How do we proceed with estimation when y is itself a probability (not 0/1)?

We invert the logit formula

Consider the binomial logit as an example:

$$P_{i1} = \frac{\exp((Z_{i1} - Z_{i2})\gamma)}{1 + \exp((Z_{i1} - Z_{i2})\gamma)}$$

$$P_{i1} = rac{\exp((Z_{i1} - Z_{i2})\gamma)}{1 + \exp((Z_{i1} - Z_{i2})\gamma)}$$

$$P_{i1} = rac{\exp\left(\left(Z_{i1} - Z_{i2}
ight)\gamma
ight)}{1 + \exp\left(\left(Z_{i1} - Z_{i2}
ight)\gamma
ight)}$$

 $1 - P_{i1} = \frac{1}{1 + \exp((Z_{i1} - Z_{i2})\gamma)}$

And the complement:

$$P_{i1} = rac{\exp\left(\left(Z_{i1} - Z_{i2}
ight)\gamma
ight)}{1 + \exp\left(\left(Z_{i1} - Z_{i2}
ight)\gamma
ight)}$$

 $\frac{P_{i1}}{1 - P_{i1}} = \exp((Z_{i1} - Z_{i2})\gamma)$

And the complement:

$$1-P_{i1}=\frac{1}{1+\exp\left(\left(Z_{i1}-Z_{i2}\right)\gamma\right)}$$

Take the ratio (denominators cancel):

$$P_{i1} = rac{\mathsf{exp}\left(\left(Z_{i1} - Z_{i2}
ight)\gamma
ight)}{1 + \mathsf{exp}\left(\left(Z_{i1} - Z_{i2}
ight)\gamma
ight)}$$

And the complement:

$$1 - P_{i1} = \frac{1}{1 + \exp((Z_{i1} - Z_{i2})\gamma)}$$

$$1 - i$$

$$\frac{P_{i1}}{P_{ii}}$$
 =

$$\frac{P_{i1}}{1 - P_{i1}} = \exp((Z_{i1} - Z_{i2})\gamma)$$

$$\frac{P_{i1}}{P_{i2}} = \epsilon$$

 $\log\left(\frac{P_{i1}}{1-P_{:1}}\right) = (Z_{i1}-Z_{i2})\gamma$

$$P_{i1} = -$$

 $P_{i1} = \frac{\exp((Z_{i1} - Z_{i2})\gamma)}{1 + \exp((Z_{i1} - Z_{i2})\gamma)}$

Start with the logit probability:

Take the ratio (denominators cancel):

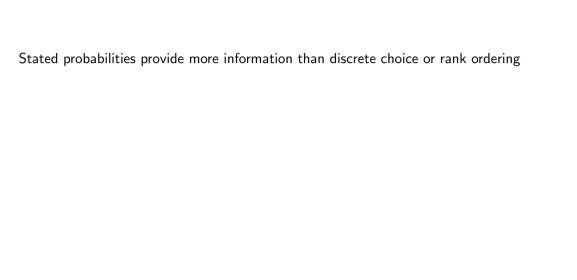
And the complement:

 $1 - P_{i1} = \frac{1}{1 + \exp\left((Z_{i1} - Z_{i2})\gamma\right)}$

 $rac{P_{i1}}{1-P_{i1}}=\exp\left(\left(Z_{i1}-Z_{i2}
ight)\gamma
ight)$

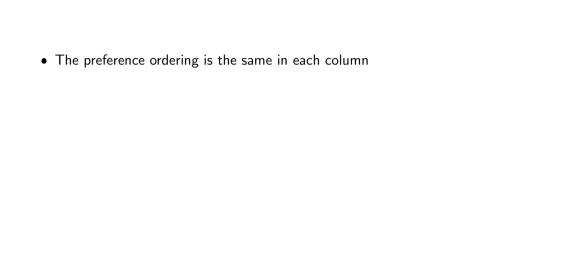
$$\log\left(\frac{P_{i1}}{1-P_{i1}}\right) = \left(Z_{i1} - Z_{i2}\right)\gamma$$

Now we're in a world where we can (in principle) use OLS to estimate γ !



Stated probabilities provide more information than discrete choice or rank ordering Consider the following responses for an individual:

Option	Discrete Choice	Rank Ordering	Stated Prob 1	Stated Prob 2	
Α	0	2	0.01	0.49	
В	1	1	0.99	0.51	
C	0	3	0.00	0.00	



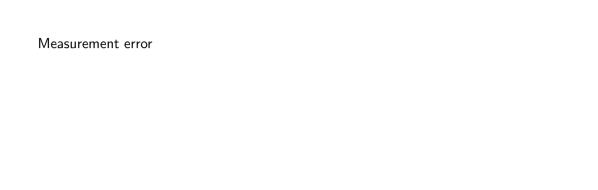
The preference ordering is the same in each column
• But the implied preference intensity is much different in the last two columns

•	The	preference	ordering	is the	same	in	each column	

Dut the involved conference intensity is much different in the last two columns

• But the implied preference intensity is much different in the last two columns

• A (0,1,0) discrete choice response corresponds to a (0,100,0) probability response



Measurement error	Measurement	error
-------------------	-------------	-------

• One concern with stated probabilities is measurement error

Measurement error

• One concern with stated probabilities is measurement error

• Rather than report 99.5%, someone may just write 100%

Measurement error

- One concern with stated probabilities is measurement error
- Rather than report 99.5%, someone may just write 100%
- But if p=0 or p=1, $\log\left(\frac{P_{i1}}{1-P_{i1}}\right)$ is undefined!

Measurement error

- One concern with stated probabilities is measurement error
- Rather than report 99.5%, someone may just write 100%
- But if p = 0 or p = 1, $\log \left(\frac{P_{i1}}{1 P_{i1}} \right)$ is undefined!
- In this case, we have to recode 0s or 1s to be small values (e.g. .001, .999)

Then, to avoid contamination, we need to use LAD instead of OLS. New equation:

 $\log\left(rac{ ilde{ ilde{ ilde{ ilde{P}}}_{i1}}{1- ilde{ ilde{ ilde{P}}}_{i1}}
ight) = \left(Z_{i1}-Z_{i2}
ight)\gamma + \eta_{i1}$

Then, to avoid contamination, we need to use LAD instead of OLS. New equation:

$$\log\left(rac{ ilde{ ilde{ ilde{P}}_{i1}}}{1- ilde{ ilde{P}}_{i1}}
ight) = \left(Z_{i1}-Z_{i2}
ight)\gamma + \eta_{i1}$$

where $ilde{P}$ is the recoded probability and η_{i1} is the difference in measurement errors

• The above is the equation for a 2-option choice set

- The above is the equation for a 2-option choice set
- With more than 2 options, we have more observations per scenario

- The above is the equation for a 2-option choice set
- With more than 2 options, we have more observations per scenario

If J = 3, we have:

$$egin{align} \log\left(rac{ ilde{\mathcal{P}}_{i1}}{ ilde{\mathcal{P}}_{i2}}
ight) &= \left(Z_{i1} - Z_{i2}
ight)\gamma + \eta_{i1} \ \log\left(rac{ ilde{\mathcal{P}}_{i3}}{ ilde{\mathcal{P}}_{i2}}
ight) &= \left(Z_{i3} - Z_{i2}
ight)\gamma + \eta_{i3} \ \end{aligned}$$

If we see N people each making $\mathcal T$ choices, then our data has $N\mathcal T(J-1)$ rows:

If we see N people each making T choices, then our data has NT(J-1) rows:

ID	Scenario	Alternative	Probability	Ζ	$Z_j - Z_B$	$\log(P_j/P_B)$
1	1	А	0.45	2.3	0.5	1.099
1	1	В	0.15	1.8	0.0	
1	1	C	0.40	3.1	1.3	0.981
1	2	А	0.25	2.5	-0.4	-0.875
1	2	В	0.60	2.9	0.0	
1	2	С	0.15	1.7	-1.2	-1.386
2	1	Α	0.10	3.2	-0.9	-1.946
2	1	В	0.70	4.1	0.0	
2	1	С	0.20	2.8	-1.3	-1.253

If we see N people each making T choices, then our data has NT(J-1) rows:

ID	Scenario	Alternative	Probability	Ζ	$Z_j - Z_B$	$\log(P_j/P_B)$
1	1	Α	0.45	2.3	0.5	1.099
1	1	В	0.15	1.8	0.0	
1	1	C	0.40	3.1	1.3	0.981
1	2	А	0.25	2.5	-0.4	-0.875
1	2	В	0.60	2.9	0.0	
1	2	С	0.15	1.7	-1.2	-1.386
2	1	А	0.10	3.2	-0.9	-1.946
2	1	В	0.70	4.1	0.0	
2	1	С	0.20	2.8	-1.3	-1.253

Drop the rows for alternative B before estimating LAD regression models