Neighborhood Choice, Information, and the Value of Amenities

Fernando Ferreira and Maisy Wong

The Wharton School

University of Pennsylvania

and NBER

July 2020

Preliminary – Please do not cite without permission

Neighborhood choice and imperfect information

- Neighborhood amenities are critical for the well-being of adults and children

  [Chay and Greenstone (2003, 2005), Bayer, Ferreira, and McMillan (2007), Bayer, Ross, Topa (2008), Linden and Rockoff (2008), Wong (2013), Chetty, Hendren, and Katz (2016), Chetty and Hendren (2018a, 2018b), Allcott et al. (2019), etc.]

- Sorting models usually assume individuals with perfect knowledge about all amenities in all neighborhoods

  [Neighborhood sorting: Bayer, Ferreira, and McMillan (2007), Wong (2013), Caetano and Maheshri (2017), Calder-Wang (2019), Almagro and Dominguez-Iino (2019), etc.]

  [Discrete choice: McFadden (1978), Berry (1997), Berry, Levinsohn, and Pakes (1995, 2004), Nevo (2001), Petrin (2002), Train (2009), Pakes and Porter (2020), etc.]

- But imperfection information likely pervasive
  - Difficult to acquire and process information about hundreds of neighborhoods and countless amenities, such as prices, demographic composition, school quality, air quality, safety, housing stock, sidewalks, etc.

This paper

1) Neighborhood choice model that allows individuals to have imperfect information about every amenity in each neighborhood

2) New neighborhood choice program that provides information about amenities to help graduating students choose where to live

3) Address unobserved neighborhood quality that varies by individual, using sufficient latent quality index (Dahl, 2002)

4) Estimates of new sorting model to recover marginal utility for rents and MWTP for amenities

Neighborhood choice model

- Utility of individual *i* choosing neighborhood *j*:

$$u_{ij} = \omega_j \beta_i^c + \varepsilon_{ij}$$

  - ω_j is a C-dimensional vector with all amenities, perfectly known by all individuals
  - β_i^c are heterogenous preferences for amenities
  - ε_{ij} is i.i.d. Type-I extreme value distribution

- Allowing for imperfect information about ω_j:

$$u_{ij} = \omega_j \beta_i^c + \Delta\omega_{ij} \beta_i^c + \varepsilon_{ij}$$

  - $\Delta\omega_{ij}$ represents individual heterogeneity on how much each person knows about every neighborhood amenity
  - Perfect information is a special case where $\Delta\omega_{ij}$ is just a null vector

Neighborhood choice model

- Econometricians only observe a limited number of amenities *k*:

$$u_{ij} = x_j\beta_i^k + \Delta\omega_{ij}\beta_i^c + \xi_j + \varepsilon_{ij}$$

  - ξ_j is the unobserved neighborhood quality index
  - Standard identification problem: $Cov(x_j, \xi_j) \neq 0$

- New identification problem: $Cov(x_j, \Delta\omega_{ij}) \neq 0$

  - Example: Some individuals may think expensive places have better schools than actual measures of quality, while other individuals may think the same about sidewalks or trees

Neighborhood choice model

- Econometricians do not observe $\Delta\omega_{ij}$ either. Collecting all unobserved terms:

$$u_{ij} = x_j\beta_i^k + \xi_{ij} + \varepsilon_{ij}$$

  - ξ_{ij} is a neighborhood-by-individual unobserved quality index
  - $Cov(x_j, \xi_{ij}) \neq 0$ seemingly intractable problem

- We will address this problem by integrating
  - Estimation of a structural model of individual neighborhood choice
  - Using individual level survey with repeated choice data, before and after receiving information about certain amenities

[Choice models and survey data: Benjamin, Heffetz, Kimball, Rees-Jones (2012), Galliani, Murphy, and Pantano (2015), Perez-Truglia and Bottan (2018), etc.]

New neighborhood choice program

- Partnership with a large professional school in the East Coast
  - Help students choose where to live upon graduation
  - School-students concerned about *cost of living* and maintaining *professional and social network*
- Pilot program rolled out in April 2019, two months prior to graduation
  - Email from Dean to introduce new neighborhood choice program
  - $25 Amazon gift card incentive, 41% take-up rate

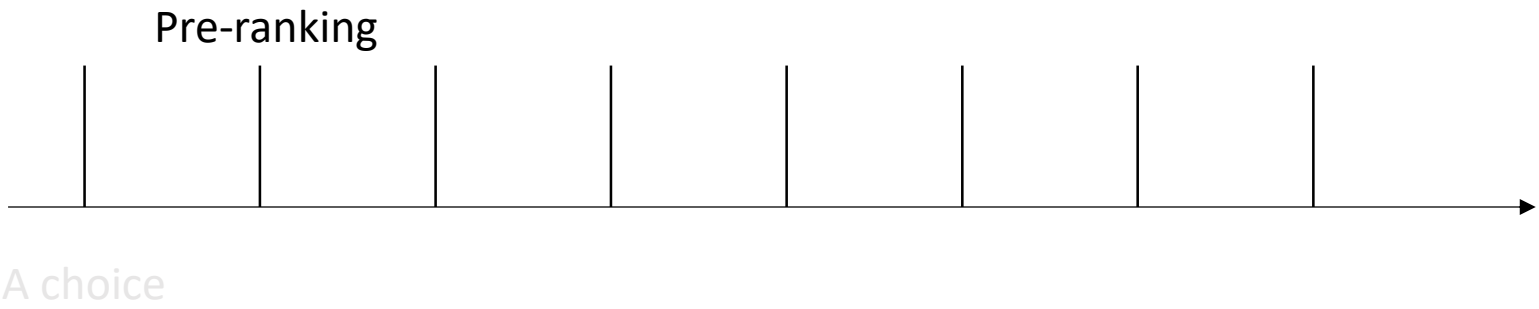
Survey timeline: Choose MSA

- Students choose MSA in which they are likely to live after graduation

MSA choice

Survey timeline: Rank neighborhoods

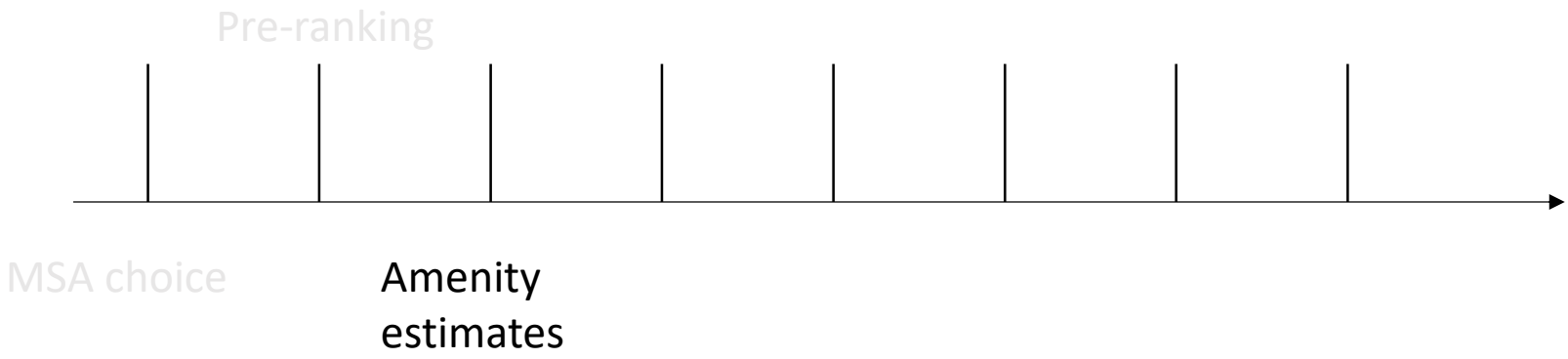
- Students rank up to 10 neighborhoods in chosen MSA
  - Presented with an average of 19 neighborhoods per MSA
  - Most students only rank 5-6 neighborhoods



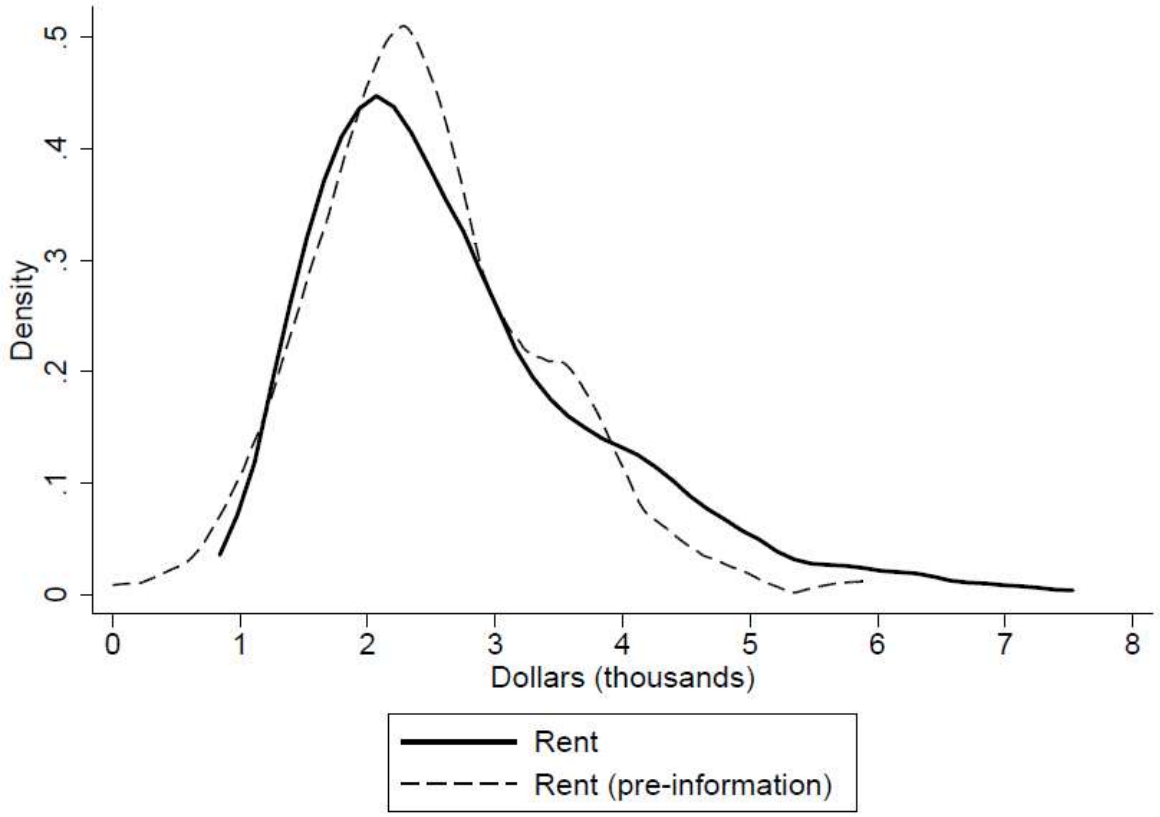
MSA	Percent	Neighborhoods	Consideration set
Atlanta	2.3	19	12
Austin	0.6	14	6
Baltimore	0.3	19	5
Boston	5.5	21	11
Chicago	4.2	22	14
DC	4.9	20	17
Dallas	1.6	18	11
Denver	1.0	19	7
Houston	0.6	15	8
LA	2.9	22	11
Miami	0.6	20	4
Minneapolis	0.3	17	3
NY	43.7	25	24
Philadelphia	6.1	20	15
San Diego	0.3	20	5
San Francisco	20.4	22	20
San Jose	1.0	18	11
Seattle	3.6	22	9
Total	100.0	353	193

Survey timeline: Amenity estimates

- Students asked to report their best estimates for:
  - Rent of average home in a neighborhood
  - Same-school network, i.e., the share of same-school peers from past cohorts in a neighborhood

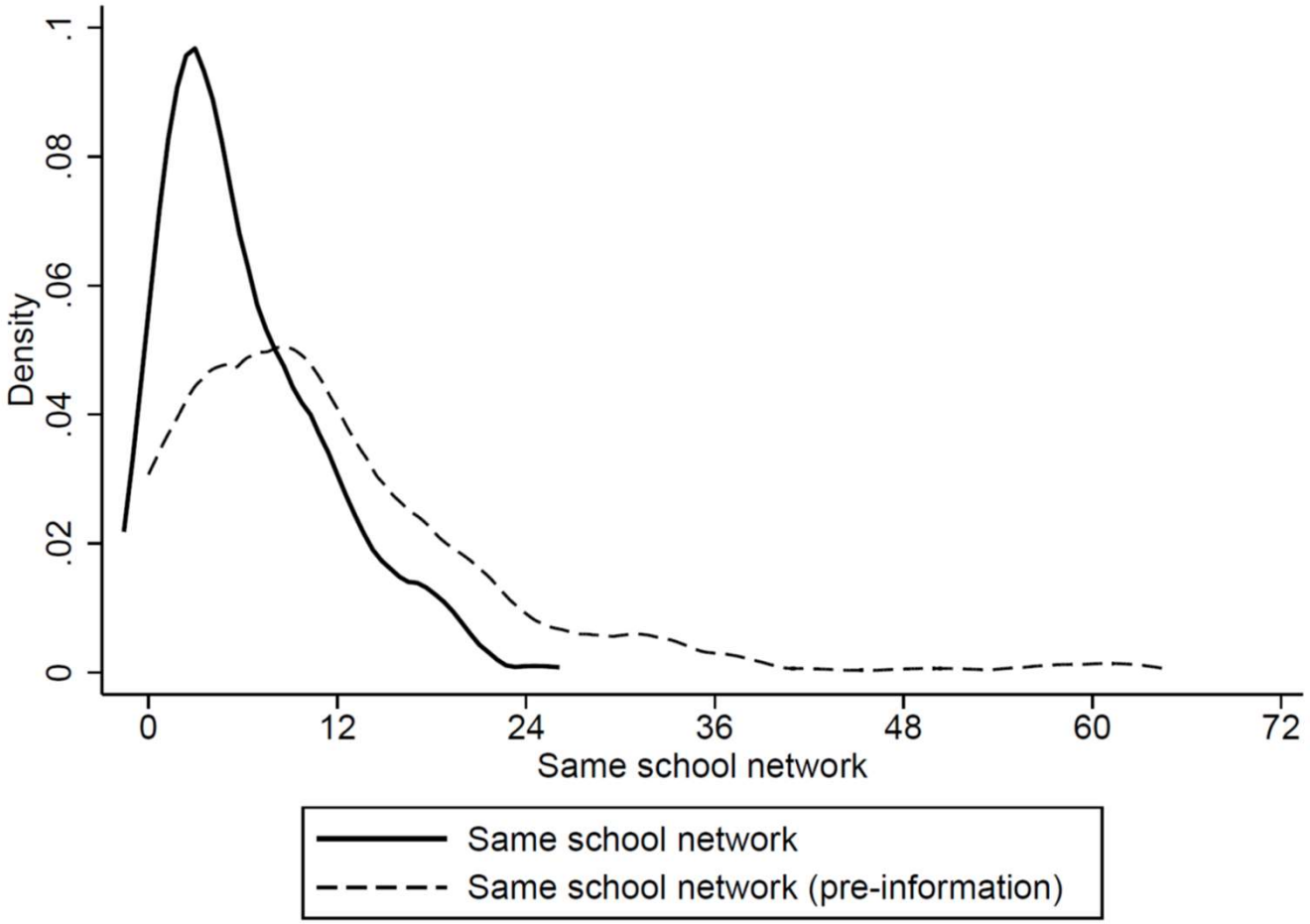


Kernel Density of Actual and Perceived Rent



- Students overestimate rents by $140 for neighborhood below $2,500
- Underestimate by $355 for neighborhoods between $2,500 and $4,000
- Underestimate by $1,200 for neighborhoods above $4,000

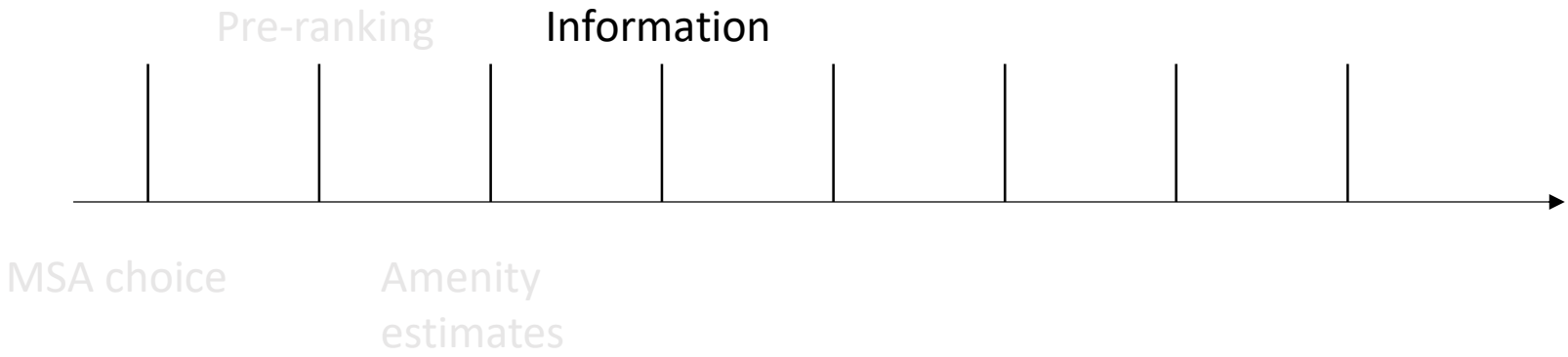
Kernel Density of Actual and Perceived Same-School Network



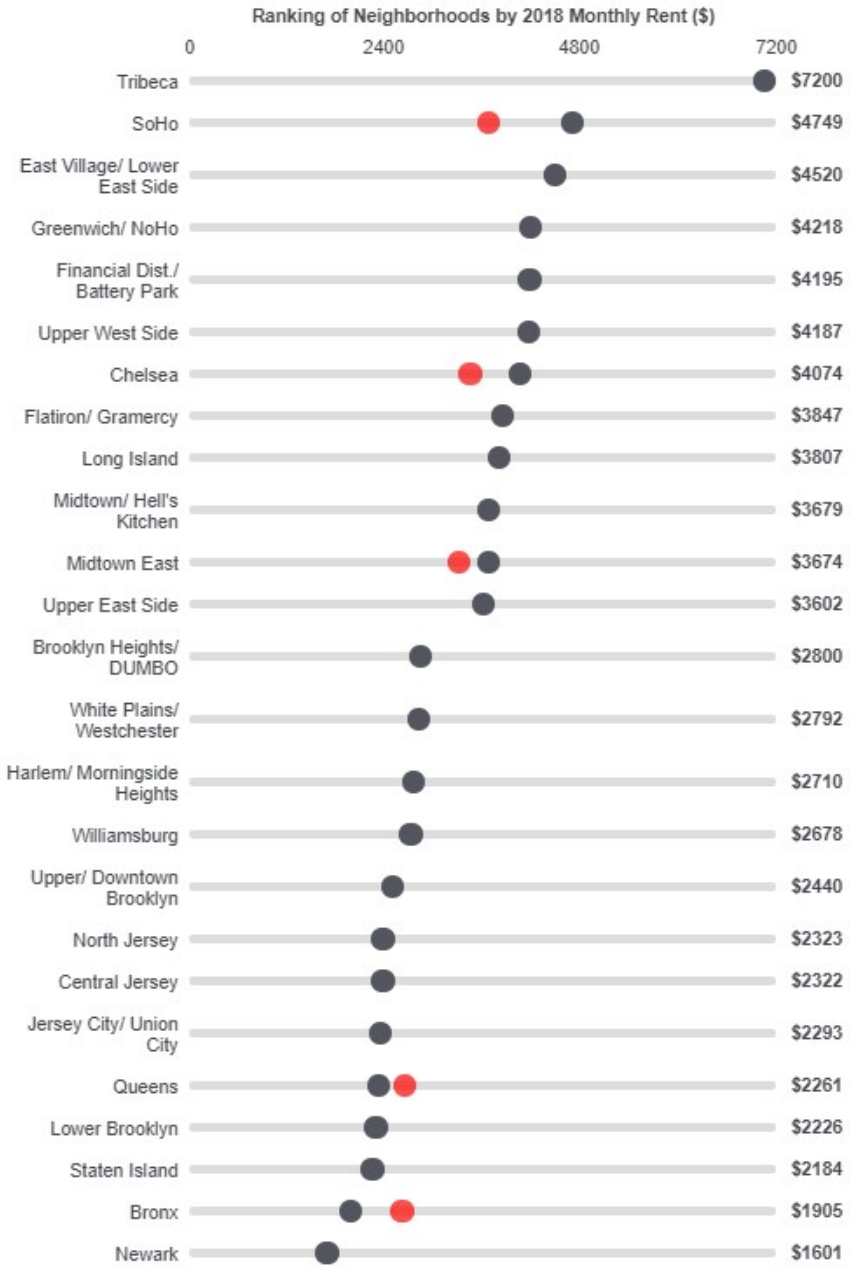
- Students overestimate same-school network by 6 percentage points

Survey timeline: Information

- Students shown information for all neighborhoods
  - Rents: Zillow rent index for average house in a neighborhood
  - Same-school network shares: Administrative information from school

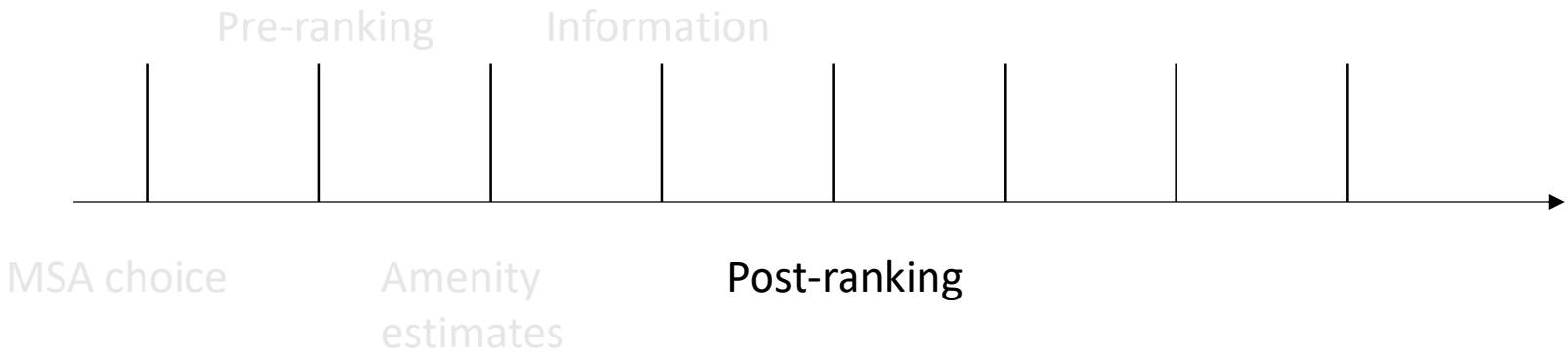


Below are the actual rents for the average home in each of the neighborhoods in New York, NY (alongside your estimate in red):



Survey timeline: Post-ranking

- Students asked to re-rank neighborhoods
  - Allowed to access new information while completing post-information ranking



Rents by pre-post top 3 rank

Panel A: Rent

		Pre	
		Yes	No
Post	Yes	Always Top 3 $3,627	Switch In $3,380
	No	Switch Out $3,844	Never Top 3 $3,007

- Always top 3 neighborhoods are more expensive than never top 3 neighborhoods – quality differences
- But students switch out of neighborhoods with higher rents, and switch into neighborhoods with lower rent
  - Imply a negative marginal utility for higher rents

Same-School Network by pre-post top 3 rank

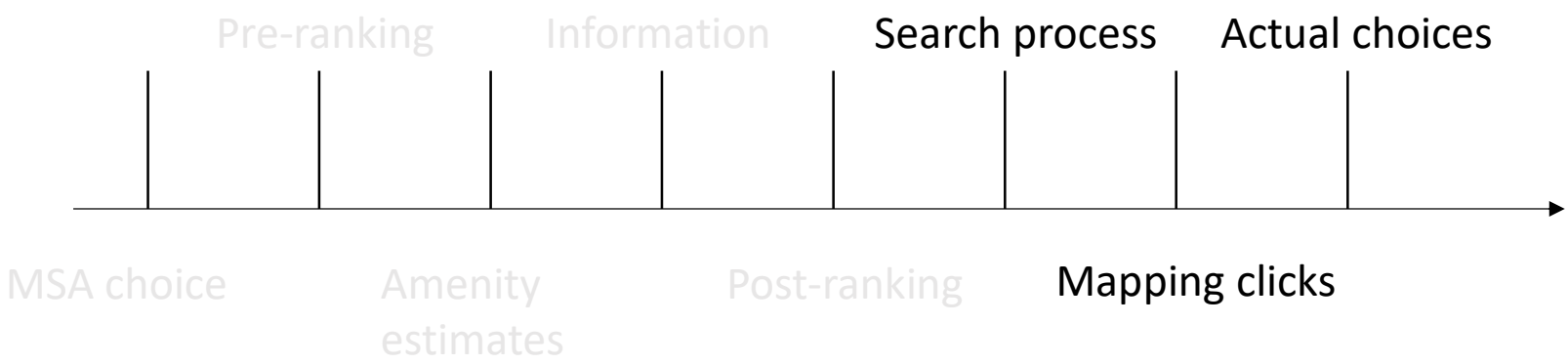
Panel B: Same-School Network

		Pre	
		Yes	No
Post	Yes	Always Top 3 7.73%	Switch In 7.18%
	No	Switch Out 5.72%	Never Top 3 3.87%

- Always top 3 neighborhoods have higher network shares than never top 3 neighborhoods
- Students switch out of neighborhoods with smaller network, and switch into neighborhoods with a larger network
  - Imply a positive marginal utility for same-school network

Survey timeline: Search and actual choices

- Students then:
  - Report the current stage of their search process
  - Access interactive maps with neighborhood information during the summer
  - Choose where to live



Estimation

- Rank order logit: Maximizes probability that individuals make the correct post-information rank ordering of neighborhoods

First stage: $$u_{ijm} = A_{jm}\beta_i^A - p_{jm}\beta_i^p + X_{jm}\beta_i + \delta_{jm} + \xi_{ijm} + \varepsilon_{ijm}$$

Second stage: $$\delta_{jm} = A_{jm}\beta_o^A - p_{jm}\beta_o^p + X_{jm}\beta_o^x + \mu_m + \eta_{jm}$$

Willingness-to-pay: $$MWTP = -\frac{\beta_o^A}{\beta_o^p} \qquad MWTP_i = -\frac{\beta_o^A + \sum z_{id}\beta_d^A}{\beta_o^p + \sum z_{id}\beta_d^p}$$

  - Amenities: Rents (*p*), same-school network (*A*), other Census features (*X*)
  - Demographics: Age, gender, married and/or with children, first-generation or minority, and citizenship status

Latent quality-index

- Our index $\tilde{\xi}_{ij}$ for neighborhood-by-individual latent quality uses the pre-intervention rankings
  - Pre-intervention rankings should capture all factors that influence individual *i*'s utility
  - Functional form: Six categorical variables that indicate if individual previously ranked neighborhood as top 1, 2, 3, 4, 5, and 6-10

- Following Dahl (2002), pre-rankings are a sufficient quality index if:
  - Individuals truthfully report pre-rankings given their best knowledge about neighborhoods
  - Changes in rankings after intervention reflect the new information about rents and same-school network

Data

	N	mean	sd
Panel A: Individuals			
Female	309	0.50	0.50
Age	309	29.36	1.78
First-gen/URM	309	0.15	0.36
Married/Kids	309	0.12	0.33
International	309	0.23	0.42
Panel B: All neighborhoods			
Same-School Network	353	5.10	5.01
Rent (in thousands)	353	2.47	1.01
Income (in thousands)	353	80.26	25.45
Bachelor's degree+	353	0.46	0.19
Non-White share	353	0.38	0.19
Panel C: Consideration set			
Same-School Network	193	7.08	5.25
Rent (in thousands)	193	2.80	1.14
Income (in thousands)	193	86.81	27.15
Bachelor's degree+	193	0.55	0.18
Non-White share	193	0.40	0.18

Table 5: MWTP Estimates With and Without Latent Quality Index

	(1)	(2)
Same-School Network	0.330***	0.328***
	(0.075)	(0.075)
Rent	-0.180	-0.314**
	(0.118)	(0.116)
Implied MWTP	**$400**	**$228****
	(257)	**(79)**
MSA FE	Y	Y
Census Characteristics	Y	Y
Demographics	Y	Y
Latent Quality Index	N	Y

- Students are willing to pay extra $228 a month in rent to live in a neighborhood with a 1 percentage point larger network
  - Average rent ~$2,500 and average same-school network share ~5%

Table 6: Heterogeneity in MWTP Estimates

Average MWTP	**$ 228*** **
	(79)
Age	-$99
	(127)
Married/Kids	-$450
	(298)
International	-$54
	(237)
First-gen/URM	$4,290
	(80,714)
Female	-$80
	(174)

Robustness

- Estimates robust to including more heterogeneity
  - Major, industry, stage of search process, previous MSA
- Including other quality indexes do not change effect of $\tilde{\xi}_{ij}$
- Student self-report why they care about neighborhoods with a larger same-school network
  - Majority cares about value of professional and social network
- Post-ranking patterns hold for mapping clicks
  - Students more likely to switch in (click in) neighborhoods with a larger network and with lower rents

Actual neighborhood choices after survey

Rents

		Pre	
		Yes	No
Post	Yes	Always Top 1 $3,977	Switch In $3,226
	No	Switch Out $3,871	Never Top 1 $3,153

Same-school network

		Pre	
		Yes	No
Post	Yes	Always Top 1 7.15%	Switch In 7.08%
	No	Switch Out 6.05%	Never Top 1 4.14%

Conclusion

- We developed a new framework for estimating preferences for neighborhood amenities under imperfect information
  - Uses new data from neighborhood choice program and exploit switchers for identification
  - Creates a new sufficient latent quality index
  - Estimates reveal a much larger price elasticity of demand, leading to higher marginal willingness to pay for amenities

- Framework can be easily adapted to other settings
  - Any amenity
  - Other decisions, such as MSA-job choices
  - Unobserved errors due to dynamics, expectations, search process, etc.

  [Bayer, McMillan, Murphy, and Timmins (2016), Caetano and Maheshri (2019)]