Discrete Choice Models

Atsushi Yamagishi

September 25, 2025

Discrete choice models

- A **discrete choice model** is a model about choices over options A, B, C, ...
- When applied to location choice, discrete choice models allow us to evaluate amenities from individual location choice data
  - In our canonical spatial economic models and Rosen-Roback models, we have used location as a unit of observations. We have not used individual-level choice data!
  - Using individual-level data allow you to incorporate individual preference heterogeneity, such as heterogeneous taste for school quality, taste for living in one's hometown etc.
- I start with basics of logit discrete choice models (Train 2009, Chapter 3).[1]
  - Multinominal logit model dates back to McFadden (1974 JPUBE)
- I then discuss applications and extensions of discrete choice models
  - Bayer, Keohane, Timmins (2009 JEEM)
  - Bayer, Ferreira, McMillan (2007 JPE)
  - Cook (2024 REStat forthcoming)

[1] This book is available at `https://eml.berkeley.edu/books/choice2.html`.

Multinominal logit: Model setup

- For individual n, location j brings the utility

$$U_{nj} = V_{nj} + \epsilon_{nj},$$

V_{nj} captures the attractiveness of location j (for individual n).

- ϵ_{nj} follows the i.i.d. type I extreme value (Gumbel) distribution. The density function and the cumulative distribution function are given by

$$f(\epsilon_{nj}) = \exp(-\epsilon_{nj}) \exp(-\exp(-\epsilon_{nj})), \quad F(\epsilon_{nj}) = \exp(-\exp(-\epsilon_{nj}))$$

- Individual n choose location j that brings the highest utility. That is, the choice probability of location i is

$$\begin{aligned} P_{ni} &= P(V_{ni} + \epsilon_{ni} > V_{nj} + \epsilon_{nj} \ \forall j \neq i) \\ &= P(\epsilon_{nj} < \epsilon_{ni} + V_{ni} - V_{nj} \ \forall j \neq i) \end{aligned}$$

Multinominal logit: Choice probability

- Conditional on specific value of ϵ_{ni}, this choice probability is written as

$$P_{ni}|\epsilon_{ni} = \Pi_{j\neq i}\exp(-\exp(-(\epsilon_{ni} + V_{ni} - V_{nj})))$$

- To derive the unconditional probability, we integrate it over the distribution of ϵ_{ni}:

$$P_{ni} = \int \underbrace{(\Pi_{j\neq i}\exp(-\exp(-(\epsilon_{ni} + V_{ni} - V_{nj}))))}_{P_{ni}|\epsilon_{ni}} f(\epsilon_{nj})d\epsilon_{ni}$$

- Calculating this integral (see Section 3.10 of Train 2009 for details), we get

$$P_{ni} = \frac{\exp(V_{ni})}{\sum_j \exp(V_{nj})}$$

  Note that the scale of V_{nj} does not matter for choice probability. We thus normalize $\exp(V_{n0}) = 1$.

- This is a simple expression for a complexity of muitinominal choice problem! A great advantage of the multinominal logic model.

Multinominal logit: Linear specification and ML estimation

- To illustrate the most common linear case, suppose that $V_{nj} = \beta X_{nj}$, where X_{nj} contains characteristics of individual n (e.g., age, gender) and characteristics of location j (e.g., amenities and wage levels).
- How can we estimate β?
  - β is the effect of having amenity X_{nj} on the utility. While it clearly relates to the "value" of amenities, we come back to the interpretation of β.
- A straightforward way is to use the choice probabilities P_{nj} in the maximum likelihood (ML)
- In $P_{ni} = \frac{\exp(\beta X_{nj})}{\sum_j \exp(\beta X_{nj})}$, choose β to best approximate the actual choice probabilities P_{nj} in the data.
  - See Train (2009, Section 3.7) for more details.
- No need for coding up yourself: `mlogit` package is available both in `R` and `STATA`.

Note: The absolute level of β and the variance of the error term

- We assume the variance of ϵ_{nj} to be $\pi^2/6$
- This is actually without loss of generality: If ϵ^*_{nj} has the variance $\sigma^2\pi^2/6$, then we can apply our model to $U_{nj} = V_{nj}/\sigma + \epsilon_{nj}$, where $\epsilon_{nj} = \epsilon^*_{nj}/\sigma$.
- Since ϵ_{ni} has the variance $\pi^2/6$, the choice probability becomes

$$P_{ni} = \frac{\exp\left((\beta^*/\sigma)X_{ni}\right)}{\sum_j \exp\left((\beta^*/\sigma)X_{nj}\right)},$$

  which is pretty much the same as before once $\beta \equiv \beta^*/\sigma$.
- That is, only β^*/σ is identified: The data do not allow you to pin down σ
  - When σ gets double, we can also double β^* to replicate the same choice probabilities.
  - Therefore, we cannot separately pin down σ and β^* just by observing choice probabilities.
- We thus focus on estimating β by normalizing the variance of ϵ_{nj} to $\pi^2/6$. See Train (2009, Section 3.2) for more discussions.

Note: Converting β into monetary units

- Both the hedonic approach and multinominal logit approach use the linear regression to uncover the value of amenities.
- However, the interpretation of β in the multinominal logit model requires caution because "β" has different meanings in two approaches.
- To see this point, recall that the canonical spatial model implies the following hedonic regression:

$$r_i = \beta X_i,$$

  where X_i includes income and amenities. Since the left-hand-side is the land price, β is in monetary units.
- In contrast, in the multinominal logit approach

$$V_{ni} = \beta X_{ni},$$

  so that β is in utility units.

Note: Converting β into monetary units

- First, to interpret β in monetary units in multinominal logit model, suppose that X_i includes 'income"

$$V_{ni} = \beta_X X_i + \beta_w w_i + \epsilon_{ni},$$

where w_i is the income level of location i. For example, X_i represents air quality (amenity).

- Then, 1 unit increase of air quality increases utility by β_X.
- This benefit of air quality improvement is equivalent to receiving β_X / β_w amount of income.
  - β_X / β_w is in monetary unit!
  - This is interpretable as (marginal) willingness-to-pay for the improvement of amenity X_i.

Expected utility

- A benefit of multinominal logit model is that we can express the expected utility in a simple way
  - This is *ex ante* utility evaluated prior to choosing the actual location after seeing realization of ϵ_{nj}.
- The expected utility (before the realization of ϵ_{nj}) is written as the following log-sum formula:

$$E(\max_j U_{nj}) = \ln\left(\sum_j \exp(V_{nj})\right) + C,$$

  where C is a constant.
- When the utility is linear in income (i.e., $\beta_w w_j$ is included in V_{nj}), we can convert this into monetary units by dividing this by β_w.
- Note that $\left(\sum_j \exp(V_{nj})\right)$ is the denominator of the logit choice probability
  - Just a coincidence, but sometimes this property is useful
- See also
  - Train (2009, Section 3.5)
  - Susumu Sato's lecture note: `https://drive.google.com/file/d/14hnb04Kn_B5FJAzjojGVYlimcxWjfeuZ/view`

Independence of Irrelevant Alternatives (IIA) Property

- The multinominal logit model substantially simplifies the choice problem, but its assumption imposes some strong structure on choice probabilities.
- The relative probability of choosing i is

$$\frac{P_{ni}}{P_{nj}} = \frac{\exp(V_{ni})}{\exp(V_{nj})}$$

- IIA property: The relative choice probability of i and j is independent of the characteristics of other choice $k \neq i, j$.
- Some extensions of discrete choice models, such as a nested logit model, do not impose the IIA property. I briefly come back to this later.

Independence of Irrelevant Alternatives (IIA) Property

- Why is the IIA a potential problem? Condition on two choices: Tokyo, and Sapporo. Let's say $P_{n,Tokyo}/P_{n,Sapporo} = 1$.
- Now consider there is a new alternative location Saitama. Then, probably $P_{n,Tokyo}/P_{n,Sapporo} < 1$ because Saitama and Tokyo are more substitutable, and many people who previously chose Tokyo now choose Saitama.
  - But this violates the IIA property.
  - a.k.a., "the red-bus-blus-bus problem."
- The IIA might not be so bad property. How bad it is just depends on the purpose of your analysis.
- But you should be aware that you are implicitly imposing some structure on the choice probabilities by using the multinominal logit model.
  - More generally, be careful about what you are implicitly assuming when you commit to a specific model.

Multinominal logit in a location choice model

- We now embed the multinominal logit model in a standard location choice model
  - I follow the formulation of Bayer et al. (2009 JEEM)
  - See Liang et al. (2024 IER) for an extension with observable types (e.g., men and women)

- Consider worker n with the following utility function:

$$U_{nj} = C_n^{\beta_C} H_n^{\beta_H} \underbrace{X_j^{\beta_X} e^{M_{n,j}+\xi_j+\epsilon_{nj}}}_{\text{Treated as constant by workers after choosing location}},$$

  - C_n is the numeraire goods consumption and H_n is the housing consumption
  - X_j is the index of location j's characteristics (amenities)
  - ξ_j is the unobserved location j's characteristics
  - $M_{n,j}$ is the mobility cost of choosing location j for individual n
  - ϵ_{nj} is the Gumbel shock.

- Worker n in location j chooses C_n and H_n to maximize U_{nj} under the budget constraint $C_n + \rho_j H_n = I_{n,j}$, yielding the optimal numeraire goods and housing consumption:

$$C_n = \frac{\beta_C}{\beta_C + \beta_H} I_{n,j}, \quad H_n = \frac{\beta_H}{\beta_C + \beta_H} \frac{I_{n,j}}{\rho_j}$$

Multinominal logit in a location choice model

- Then, the log (indirect) utility function is written as follows:

$$\beta_I \ln I_{n,j} + M_{n,j} + \underbrace{(\beta_X \ln X_j - \beta_H \ln \rho_j + \xi_j)}_{\theta_j} + \epsilon_{n,j},$$

  where θ_j summarizes the "fundamental" attractiveness of location j that is independent of individual n's characteristics.

- Workers maximize the above indirect utility in choosing their location
  - The choice probability takes the logit formula due to ϵ_{nj}, where $V_{nj} = \beta_I \ln I_{n,j} + M_{n,j} + \theta_j$.
- Two-step estimation, as in Berry, Levinson, Pakes (1995 ECMA):
  - Estimate θ_j and other parameters using the maximum likelihood.
    - This step also estimates β_I and $M_{n,j}$, which takes account of individual heterogeneity in income and moving costs.
  - Estimate the regression model $\theta_j = \beta_X \ln X_j - \beta_H \ln \rho_j + \xi_j$
    - Due to endogeneity concern, Bayer et al. used calibrated value for β_H and take the first-difference. Then use the distant plants as an IV for air quality.
    - Due to endogeneity of ρ_j, Bayer et al assumes $\beta_H = 0.2$ based on previous literature.

Estimation results: Bayer et al. (2009)

People hate to move away from home state or region.

Table 4
Results from first-stage discrete choice model of residential location decision.

Variable	Parameter	Coefficient	t-Statistic
Migration cost			
State	μ_S	–2.900	–22.0
Region	μ_{R1}	–0.855	–11.5
Macro-region	μ_{R2}	–0.591	–12.5
Marginal utility of income	β_I	0.673	48.4

Table 5
Results from second-stage regressions.

Dependent variable	OLS		IV		
$\Delta\theta$+0.25$\Delta\ln\rho$	(1)	(2)	(3)	(4)	(5)
$\Delta\ln(PM)$	–0.086 (0.060)	–0.107** (0.054)	–0.255** (0.110)	–0.286*** (0.109)	–0.230** (0.101)
$\Delta\ln(Crime)$		0.010 (0.067)		0.008 (0.073)	0.024 (0.068)
$\Delta\ln(Prop.\ tax)$		0.359* (0.186)		0.396* (0.203)	0.346* (0.188)
$\Delta\ln(Govt.\ exp.)$		0.112*** (0.039)		0.131*** (0.042)	0.114*** (0.039)
$\Delta\ln(White)$		–0.064 (0.389)		–0.132 (0.424)	–0.034 (0.394)
$\Delta\ln(Health)$		–0.001*** (0.000)		–0.001*** (0.000)	–0.001*** (0.000)
$\Delta\ln(Arts)$		0.000 (0.000)		0.000 (0.000)	0.000 (0.000)
$\Delta\ln(Transport)$		0.000 (0.000)		0.000 (0.000)	0.000 (0.000)
$\Delta\ln(Employment)$		–0.367 (0.391)		–0.612 (0.424)	–0.319 (0.397)
$\Delta\ln(Manuf.\ est.)$		0.023 (0.087)		0.271*** (0.081)	0.020 (0.088)
$\Delta\ln(Population)$		0.820*** (0.146)			0.823*** (0.148)
Constant	–0.020 (0.050)	–0.058 (0.053)	–0.087 (0.063)	–0.049 (0.065)	–0.101 (0.061)
Regional dummies	Yes	Yes	Yes	Yes	Yes
R^2	0.08	0.32	0.05	0.19	0.31
Observations	242	242	242	242	242

Notes: Standard errors in parentheses.
* Significance at 10%.
** Significance at 5%.
*** Significance at 1%.

Amenity (air quality) evaluation: Bayer et al. (2009)

Table 6
Estimated marginal willingness to pay for air quality.

Measure	Hedonics		Residential sorting			
	OLS	IV	OLS	IV		
	Full specification (1)	Full specification (2)	Full specification (3)	Full specification (4)	No covariates (5)	No control for population (6)
WTP Elasticity	0.06	0.13	0.16	0.34	0.38	0.42
MWTP ($)	25.40	55.20	69.10	148.70	164.72	184.89

Notes: Specifications (1)–(4) are full specifications. Specification (5) includes no covariates. Specification (6) includes no control for population. "Hedonics" coefficients are taken from the wage-hedonic model summarized in Table 4 (columns 2 and 4). "Residential sorting" coefficients are taken from Table 5; columns 3–6 above correspond to columns 2, 5, 3, and 4 in Table 5, respectively. Marginal willingness to pay (MWTP) is calculated by multiplying the regression coefficients by the median household income in constant 1982–1984 dollars ($15,679) and dividing by the median PM10 concentration in the sample (36.0 $\mu g/m^3$). Figures for the wage-hedonic model exclude the estimated effects of PM10 on income, which were insignificant in the IV model. All estimates are in constant 1982–1984 dollars.

- Compared to the Rosen-Roback approach, higher WTP for air quality.
- Intuition: People stay in the polluted region either because (i) they do not hate air pollution that much or (ii) they face moving costs.
  - The Rosen-Roback approach assumes (i) is the only reason, so it (misleadingly) estimates lower WTP for air quality.

Bayer et al. (2007 JPE)

- So far we have assumed that preferences for amenities are homogeneous, but *heteroegneous preferences* may exist
  - Homogeneous preferences are implied by β being indepenent of individual n's characteristics.
- *Sorting:* If different people sort into different locations, then preferences for amenities may greatly vary by location
  - Preferences for amenities may "jump" at the border, and this may affect how we should convert land price differences into amenity values.
- Bayer et al. (2007) illustrate the importance of preference heterogeneity for school quality
  - Revisiting Black (1999) we have seen before
- Two themes of this paper
  - How can we estimate the mean preferences for house characteristics (e.g., school quality) in the presence of preference heterogeneity?
  - How much does failing to account for sorting and preference heterogeneity matter in the estimation of school quality?

Sorting at the border

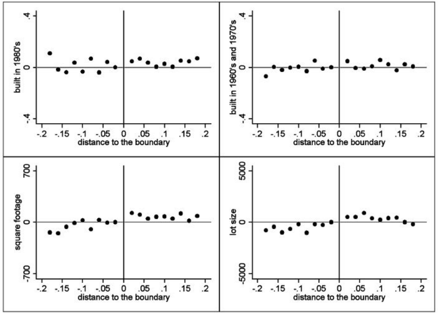


Fig. 3.—Transactions data housing characteristics around the boundary. Each panel is constructed using the following procedure: (i) regress the variable in question on boundary fixed effects and on 0.02-mile band distance to the boundary dummy variables; (ii) plot the coefficients on these distance dummies. Thus a given point in each panel represents this conditional average at a given distance to the boundary, where negative distances indicate the low test score side.

- Housing characteristics seem to be similar across the border
- This is something you should check when using a border design!

Sorting at the border

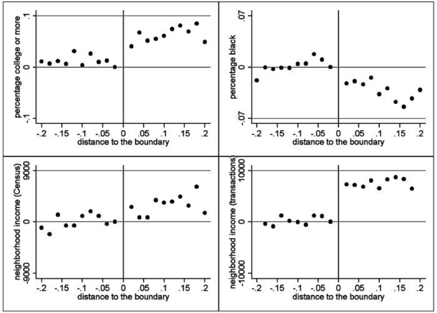


FIG. 4.—Neighborhood sociodemographics around the boundary. Each panel is constructed using the following procedure: (i) regress the variable in question on boundary fixed effects and on 0.02-mile band distance to the boundary dummy variables; (ii) plot the coefficients on these distance dummies. Thus a given point in each panel represents this conditional average at a given distance to the boundary, where negative distances indicate the low test score side.

- But we see a jump of residents' education level, race, and income at the border of school districts
- This is also something you should check when using a border design (if you can)!

Model with heterogeneous preferences for school quality

- Multinominal logit model with preference heterogeneity for school quality
  - Preference heterogeneity is a heart of Berry, Levinson, Pakes (1995). But unlike the original BLP, this paper uses individual-level microdata of location choice as in BLP (2004 RES)

- The (indirect) utility of household n for housing j is given by

$$U_j^n = \alpha_x^n x_j - \alpha_p^n p_j - \alpha_d^n d_j^n + \theta_{bj} + \xi_j + \epsilon_j^n$$

  - x_j: observable attributes of house j (including school quality)
  - p_j: price of house j
  - d_j^n: distance of house j to place of work of household n
  - θ_{bj}: the school boundary fixed effect relevant for house j
  - ξ_j: unobservable attribute of house j
  - ϵ_j^n: idiosyncratic utility of house j for household n

- Each household's marginal utility of each attribute is allowed to vary with its observable attributes z_k^n:

$$\alpha_c^n = \alpha_{0c} + \sum_{k=1}^{K} \alpha_{kc} z_k^n$$

Model with heterogeneous preferences for school quality

- Let $\delta_j \equiv \alpha_{x0}x_j - \alpha_{0p}p_h + \theta_{bj} + \xi_j$ be the "baseline utility" of house j that is common to all households.
- Let $\lambda_j^n \equiv (\sum_{k=1}^{K} \alpha_{kx} z_k^n)x_j - (\sum_{k=1}^{K} \alpha_{kp} z_k^n)p_j - (\sum_{k=1}^{K} \alpha_{kd} z_k^n)d_j^n$, which is location j's attractiveness that varies across individual's observable characteristics z.
  - Since commuting cost depends on individual n's workplace, it appears in λ_j^n but not in δ_j
- Then, $U_j^n = \delta_j + \lambda_j^n + \epsilon_j^n$.
- Two-step estimation:
  - Estimate δ_j and parameters in λ_j^n by the Maximum Likelihood of standard multinominal logit.
  - Using the estimated δ_j, use the linear regression for estimating parameters in $\delta_j = \alpha_0 x_j - \alpha_{0p}p_h + \theta_{bj} + \xi_j$.

Relationship to the standard hedonic regression

- Rearranging $\delta_j = \alpha_0 x_j - \alpha_{0p} p_h + \theta_{bj} + \xi_j$, we get

$$P_j + \frac{1}{\alpha_{0p}}\delta_j = \frac{\alpha_{0x}}{\alpha_{0p}} x_j + \frac{1}{\alpha_{0p}}\theta_{bj} + \frac{1}{\alpha_{0p}}\xi_j$$

- This is a regression equation that allows us to estimate the mean preferences for housing attributes x_j
- This looks almost like a standard hedonic regression (recall Black 1999), but the left-hand-side has the "adjustment term" $\frac{1}{\alpha_{0p}}\delta_j$.
  - To estimate the mean willingness-to-pay, we need to "adjust the price upward" when the option j is popular, and vice versa.
- But note that Bayer et al (2007) approach also allows us to estimate the preference heterogeneity, not just the mean preferences.

Results: Sorting and preference heterogeneity matters

- The school quality effect drops by including boundary fixed effects (Black 1999)
- But it substantially drops after including socioeconomic characteristics of school districts.
- Evidence of sorting according to school districts

TABLE 7
DELTA REGRESSIONS: IMPLIED MEAN WILLINGNESS TO PAY
SAMPLE: WITHIN 0.20 MILE OF BOUNDARY (N = 27,458)

Boundary fixed effects included	No	Yes
	A. Excluding Neighborhood Sociodemographic Characteristics	
	(1)	(2)
Average test score (in standard deviations)	97.3	40.8
	(14.0)	(5.5)
	B. Including Neighborhood Sociodemographic Characteristics	
	(3)	(4)
Average test score (in standard deviations)	18.0	19.7
	(8.3)	(7.4)
% block group black	−404.8	−104.8
	(41.4)	(36.9)
% census block group Hispanic	−88.4	−3.5
% block group with college degree or more	183.5	104.6
	(26.4)	(31.8)
Average block group income (/10,000)	30.7	36.3
	(3.7)	(6.6)

NOTE.—All regressions shown in the table also include controls for whether the house is owner-occupied, the number of rooms, year built (1980s, 1960–79, pre-1960), elevation, population density, crime, and land use (% industrial, % residential, % commercial, % open space, % other) in 1-, 2-, and 3-mile rings around each location. The dependent variable is the monthly user cost of housing, which equals monthly rent for renter-occupied units and a monthly user cost for owner-occupied housing, calculated as described in the text. Standard errors corrected for clustering at the school level are reported in parentheses.

Results: Sorting and preference heterogeneity matters

- There are substantial preference heterogeneity for school quality and other neighborhood characteristics
- This is behind the substantial drop of the school quality value after controlling for socioeconomic characteristics
  - People living in "good" school districts are willing to pay more for better school quality, bidding up the housing prices of good school districts and yielding larger housing price discontinuity

TABLE 8
HETEROGENEITY IN MARGINAL WILLINGNESS TO PAY FOR AVERAGE TEST SCORE AND NEIGHBORHOOD SOCIODEMOGRAPHIC CHARACTERISTICS

	Average Test Score +1 SD	Neighborhood Sociodemographics: +10% Black vs. White	+10% College-Educated	Block Group Average Income +$10,000
Mean MWTP	19.69	−10.50	10.46	36.3
	(7.41)	(3.69)	(3.18)	(6.60)
Household income (+$10,000)	1.38	−1.23	1.41	.86
	(.33)	(.37)	(.21)	(.12)
Children under 18 vs. no children	7.41	11.86	−16.07	2.37
	(3.58)	(3.03)	(2.25)	(1.17)
Black vs. white	−14.31	98.34	18.45	−1.16
	(7.36)	(3.93)	(4.52)	(2.24)
College degree or more vs. some college or less	13.03	9.19	58.05	.31
	(3.57)	(3.14)	(2.33)	(1.40)

NOTE.—The first row of the table reports the mean marginal willingness to pay for the change reported in the column heading. The remaining rows report the difference in willingness to pay associated with the change listed in the row heading, holding all other factors equal. The full heterogeneous choice model includes 135 interactions between nine household characteristics and 15 housing and neighborhood characteristics. The included household characteristics are household income, the presence of children under 18, and the race/ethnicity (Asian, black, Hispanic, white), educational attainment (some college, college degree or more), work status, and age of the household head. The housing and neighborhood characteristics are the monthly user cost of housing, distance to work, average test score, whether the house is owner-occupied, number of rooms, year built (1980s, 1960–79, pre-1960), elevation, population density, crime, and the racial composition (% Asian, % black, % Hispanic, % white) and average education (% college degree) and household income for the corresponding census block group. Standard errors are reported in parentheses.

Cook (2024 REStat forthcoming)

- So far, we have assumed that each location choice j corresponds to residential choice.
- Cook (2024) models visits to amenity facilities in a city (restaurants, parks etc) using the discrete choice model.
  - Where to eat out today?
  - Which park do you go to?
- He uses travel data within a city to evaluate these amenity facilities
  - Smartphone GPS data that closely tracks people's movement
  - See Miyauchi, Nakajima, Redding (2025 QJE) and Arai et al. (2023 wp) for examples of studies using smartphone data in Japan.
- Intuition: when more people visit facility j, this facility j should be attractive.
- Using this, we can assess how amenities access varies by location.
  - For instance, who has barriers to access to supermarkets? The food desert problem (c.f., Allcott et al. 2019 QJE).

Cook (2024 REStat forthcoming)

- Let θ_j be the attractiveness of amenity facility j.
- The utility of visiting j for individual n is

$$U_{nj} = \theta_j - \kappa^H d_{nj}^H - \kappa^W d_{nj}^W + \epsilon_{nj},$$

  where d_{nj}^H (d_{nj}^W) is the distance from individual n's home (workplace) to facility j
- The expected utility may be used as an "amenity quality index" for individual n:

$$\ln\left(\sum_{j \in \text{All available amenities around individual } n} \exp(\theta_j - \kappa^H d_{nj}^H - \kappa^W d_{nj}^W)\right)$$

- This amenity quality index varies by location of individual n.
  - If we do this separately for different types of individuals (e.g., gender), then we get amenity quality index for different types of people.

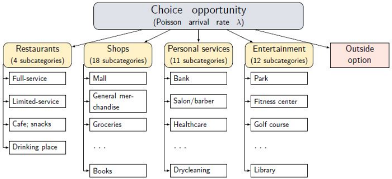
Cook (2024 REStat forthcoming)

- Cook considers a slightly more complicated situation because he suspects that $\epsilon_{n,j}$ may not be i.i.d. so that similar facilities have similar ϵ.

- He uses a nested logit model, in which ϵ is allowed to be correlated within a category of facilities.
  - This relaxes the IIA: If McDonalds is excluded from the choice, those who previously going to McDonalds tend to go to BurgerKing (a choice in the same "fast food" nest) rather than Japanese restaurants because McDonalds would have a similar idiosyncratic term to BurgerKing
  - In a standard logit model, this effect is absent because ϵ is iid across all options

- Things are a bit more complicated but many (but not all) properties of multinominal logit are preserved. See Train (2009, Chapter 4) for more details.

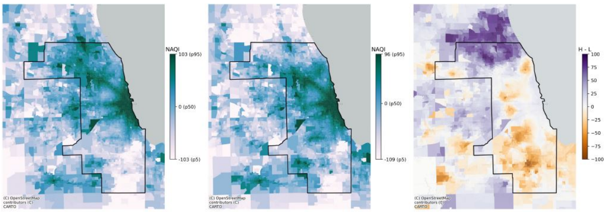
Figure 2: Model structure



Cook (2024 REStat forthcoming)

Figure 5: NAQI: Chicago

(a) Higher income (b) Lower income (c) Difference (H - L)



Note: This figure illustrates the estimated block group level NAQI values in Equation 4.9 for the Chicago-Naperville-Elgin, IL-IN-WI CBSA, zoomed in to Cook County (outlined in black). NAQI units correspond to minutes of weekly driving time relative to the median neighborhood in a CBSA.

- Using the GPS data from the US to construct the amenity quality index, Cook finds that high-income and low-income people agree upon attractive locations
- Horizontal taste differences do not account for income sorting.
    - But better places tend to be occupied by the rich people because they are willing to pay more rents.

Taking stock

- Discrete choice model (multinominal logit model) is a powerful tool for evaluating amenities
- My take is that discrete choice model is preferred when (i) you have individual-level choice data and (ii) you are interested in preference heterogeneity
  - You can incorporate individual circumstances, such as mobility costs (Bayer et al. 2007; Bayer et al. 2009).
  - Estimating a discrete choice model without individual-level data is also possible, but more complicated (Berry et al. 1995)
  - In contrast, the canonical spatial model and the Rosen-Roback models provide sharp simple empirical implications in this context
- When there is no preference heterogeneity, both discrete choice approach and the Rosen-Roback type hedonics are convenient.
- In addition, multinominal logit model is an important building block for various spatial models
  - We have seen the first example of this in Bayer et al. (2009).
  - We will see more in discussing the quantitative spatial economic (QSE) models.