Introduction and static discrete choice models

Olivier De Groote

TSE

MRes EEE, 2021-2022

Outline

Introduction I

- Economics is the science which studies human behavior as a relationship between ends and scarce means which have alternative uses (Robbins, 1932)
- Choices transform these scarce means into outcomes, hoping this will increase utility
- Many of these choices are discrete
  - Mode of transportation
  - Type of car
  - School or study program
  - ...

Introduction II

- Discrete choice models applied to choice data help us to learn why agents make certain choices
  - Derive price elasticity of demand (own and cross)
  - Derive willingness to pay for different characteristics of goods
  - ...
- Once we understand why they make the choices they do, we can use the models to make counterfactual predictions of choices
  - Changes in prices (taxes)
  - Introduction of a new product
  - Constraints on choice sets

Introduction III

- Why is this interesting?
  - "Science" argument: it allows us to isolate the impact of different channels in explaining the world
  - "Policy" argument: we can predict (quantitatively) the impact of a new policy
- We can also calculate the impact on welfare of these counterfactuals
  - Thought experiments: who gains/suffers from what?
  - Predicting new policies: will there be gains and for whom?
  - Evaluation of policies: were there gains and for whom?

Structural vs reduced form

- The types of model we consider are called “structural”
- First step: estimate the “primitives” of a model
  - Utility, profit, productivity,....
- Ultimate goal: use them for counterfactual simulations where primitives are fixed but the situation changes
  - Subsidy, choice restriction, “shutting down” a channel,...

Structural vs reduced form

- In contrast to "reduced form" methods (DID, IV, RDD...) need to be more explicit about primitives
- Big difference in practice: even without a story behind your identification strategy, you can get some result out of it... this can (and should!) make people very suspicious about results
- An intuitive approach to discuss identification is therefore always the minimum required
  - Keane (2010): "What are the key features of the data, or the key sources of (assumed) exogenous variation in the data, or the key a priori theoretical or statistical assumptions imposed in the estimation, that drive the quantitative values of the parameter estimates, and strongly influence the substantive conclusions drawn from the estimation exercise?"

Structural vs reduced form

- A formal way of discussing identification can be found in the econometrics literature on the subject, see in particular Magnac & Thesmar (2002), Hu & Shum (2012), Abbring & Daljord (2020) in the context of applications we will discuss
- These papers focus on nonparametric identification of models, i.e. (to what extent) do you need the parametric structure you impose to be able to estimate primitives
- Keep in mind that this is not a necessary, nor sufficient condition to obtain reliable results
  - Kalouptsidi, Scott, Souza-Rodrigues (2021) show identification of counterfactuals in DDC, despite the model being unidentified
  - Rust (2010) argues that many "parametric" restrictions (e.g. lognormal distribution of wages) might be weaker than several "nonparametric" restrictions (e.g. rational expectations)

Structural vs reduced form

- Reduced form is not without danger either
  - Usually clear about identification, but not always about the theoretical framework -> do not make mistake in thinking it doesn't need one
  - See also Belzil (2007) and Keane (2010) who show that many papers make strong (but implicit!) assumptions that could be relaxed in a structural approach (in particular: static optimization, or homogeneous treatment effects)
- Key messages when you do empirical work:
  - Understand why the data are telling you something: is it the result of an arbitrary assumption you made? Or is it really the answer to your question?

Discrete choice models: how to proceed

1. Specify a choice set among which agent i chooses an option j
2. Specify a utility function for agent i if she chooses option j
3. Assume agents choose the option j that maximizes their utility
4. Estimate unknown parameters of the utility function
   - 4.1 Individual data: derive a likelihood function using ML
   - 4.2 Aggregate data: derive a regression equation (OLS or 2SLS) on market share information
   - 4.3 Or (both cases): derive moment conditions that have to be satisfied in the data to estimate using GMM

Static vs dynamic discrete choice models

- First part of this course: static discrete choice
  - Assumption: agents optimize utility today
- Second part of this course: dynamic discrete choice
  - Assumption: agents optimize expected lifetime utility
- One can think of agents to always be dynamic, however to characterize their choices, a static model can be fine in many circumstances
- One of the goals of this course is to make you understand when dynamics are (not) crucial, never try to model everything

Goal of this course I

- Rust (2010) writes "... What explains this state of affairs? Keane identifies one important reason: structural econometrics is difficult. [...] structural models rarely admit a convenient closed form solution that can be estimated via regressions or other standard statistical routines in popular statistical packages such as Stata"
- One of the goals of this course is to show that this is no longer true! What we will do is feasible in standard software.
- Discrete choice models (and structural models more generally) often rely on a large number of untestable assumption
- These assumptions result from
  - Data limitations (e.g. a low number of observations, unobserved heterogeneity,...)
  - Computational limitations

Goal of this course II

- Some of these assumptions seem ad hoc but help estimation a lot
- Contributions on this front are crucial to move forward as a science
  - Time constraints: you don't have the model ready for your JMP now, so it cannot run for several years...
  - Fixing mistakes in code? Imagine only being able to fix a mistake once a.. month, year..
  - Do you trust an estimator you only ran once?
  - Improving model fit?
  - Reliable standard errors without resampling procedures?
  - Robustness checks?

Goal of this course III

- Applied economists have to solve a trade-off, we need a model that is not too complex but still flexible enough to capture what's of first order importance
- Before diving into models, always ask yourself first (and convince the audience of your answers!):
  - Do we need a structural model?
  - Do we need a dynamic model?
  - Do we need to fully describe the model?
  - Do we need to model ... as a choice variable, an equilibrium outcome or can it be exogenous?
  - Can I make the model compatible with a fast estimation strategy?

Goal of this course IV

- The aim of this course is to (1) discuss models that include some or all of the above characteristics, and (2) discuss recent applications of these methods. Both should help you to answer the above and pursue your own research using these methods.

- Focus on applied work in the context of environmental and education (labor) economics

- Details about econometrics is left for own reading and/or other courses

Some background I

- Main references in this course are for applied economists
  - Static discrete choice: Train, K. Discrete Choice Methods with Simulation. Cambridge; New York: Cambridge University Press, 2009.
  - Dynamic discrete choice and CCP: Arcidiacono, P., and Ellickson, P. "Practical Methods for Estimation of Dynamic Discrete Choice Models." Annual Review of Economics 3 (2011): 363–94.
  - IV estimation of dynamic models and the ECCP estimator: Kalouptsidi, M., Scott, P.T., Souza-Rodrigues, E., 2019. "Linear IV Regression Estimators for Structural Dynamic Discrete Choice Models." Journal of Econometrics 222(1): 778-804.

Some background II

- Other (larger) reviews

    - Aguirregabiria, V., and Mira, P. "Dynamic Discrete Choice Structural Models: A Survey." Journal of Econometrics 156, no. 1 (2010): 38–67.

    - Keane, M. Todd, P. and Wolpin, K. "The Structural Estimation of Behavioral Models: Discrete Choice Dynamic Programming Methods and Applications." In Handbook of Labor Economics, 4:331–461. Elsevier, 2011.

Some background III

- Some key contributions in dynamic discrete choice
  - Rust, J. "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." Econometrica 55, no. 5 (1987): 999–1033.
  - Hotz, J., and Miller, R. "Conditional Choice Probabilities and the Estimation of Dynamic Models." The Review of Economic Studies 60, no. 3 (1993): 497–529.
  - Keane, M., and Wolpin, K. "The Career Decisions of Young Men." Journal of Political Economy 105, no. 3 (1997): 473–522.
  - Aguirregabiria, V., and Mira, P. "Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models." Econometrica 70, no. 4 (2002): 1519–1543.
  - Arcidiacono, P., and Miller, R. "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity." Econometrica 79, no. 6 (2011).

Applications I

- Make sure you read these paper before class
- Common theme: educational choices and green technology adoption
  - Applications where dynamics are often crucial
- Static discrete choice
  - Abdulkadiroğlu, A., Agarwal, N. and Pathak, P. "The Welfare Effects of Coordinated Assignment: Evidence from the New York City High School Match." American Economic Review 107, no. 12 (2017): 3635–89.
  - Grigolon, L., Reynaert, M. and Verboven, F. "Consumer Valuation of Fuel Costs and Tax Policy: Evidence from the European Car Market." American Economic Journal: Economic Policy 10, no. 3 (2018): 193–225.

Applications II

- Dynamic discrete choice
  - Arcidiacono, P. "Affirmative Action in Higher Education: How Do Admission and Financial Aid Rules Affect Future Earnings?" Econometrica 73, no. 5 (2005): 1477–1524.
  - De Groote, O. and Verboven, F. "Subsidies and Time Discounting in New Technology Adoption: Evidence from Solar Photovoltaic Systems." American Economic Review 109, no. 6 (2019): 2137-2172.
  - Heckman, J.J., Humphries, J.E., Veramendi, G. Returns to Education: The Causal Effects of Education on Earnings, Health, and Smoking. Journal of Political Economy 126 (2018).

Applications

- Other recommended reading of recent applications in dynamic discrete choice (will also refer to them later):
  - Housing using CCP: Murphy (2018)
  - Education using CCP: Arcidiacono et al. (2016), Declercq & Verboven (2018), De Groote (2020)
  - Labor/trade using ECCP: Traiberman (2019)
  - Housing using ECCP: Almagro & Domınguez-Iino (2021)
  - Agricultural using ECCP: Scott (2013)
  - Causal returns schooling and experience without fully specified model: Ashworth et al. (2020)

Evaluation

- Participation in class
- Coding assignment
  - Individual or group of 2
  - Simulated data of educational choices
  - Estimate a static and dynamic model and compare counterfactuals
  - Not a paper, answer straight to the question and keep it short
  - Deadline: 7 October 2021
- Paper assignment
  - Motivate RQ for which you should use a dynamic discrete choice model
  - Describe dataset
  - Propose model and estimation strategy
  - 5 to 10 pages
  - Deadline: 17 December 2021

The utility function I

- Let $u_{ij} + \epsilon_{ij}$ be the utility agent i derives from choosing alternative j
  - $u(X_j, S_i)$ is the deterministic component, or representative utility
    - X_j is a vector of product characteristics
    - S_i is a vector agent characteristics
  - ϵ_{ij} is a random taste shock
    - uncorrelated with X_j and S_i
    - uncorrelated over alternatives
- If one assumes an iid extreme value type 1 distribution on ϵ_{ij}, it leads to two types of "logit" models

The utility function II

- First type: multinomial logit:

$$u_{ij} + \epsilon_{ij} = S_i' \alpha_j + \epsilon_{ij}$$

  - with α_j a vector of parameters that describes how each agent's characteristic influences the utility for good j
  - example: the utility of a school

$$u_{ij} + \epsilon_{ij} = \alpha_j^0 + \alpha_j^{IQ} IQ_i + \epsilon_{ij}$$

The utility function III

- Second type: conditional logit

$$u_{ij} + \epsilon_{ij} = X_j'\eta + \epsilon_{ij}$$

  - with η a vector of parameters that describes how each characteristic of good j influences utility

  - example: the utility of a car

$$u_{ij} + \epsilon_{ij} = \eta^0 + \eta^{MPG} MPG_j + \eta^P P_j + \epsilon_{ij}$$

The utility function IV

- These can be seen as special cases of a more general model

$$u_{ij} + \epsilon_{ij} = S_i' \omega X_j + \epsilon_{ij}$$

  with ω a matrix of parameters that captures all interactions between agent and product characteristics

- multinomial logit: use dummy variables for each j in X_j
- conditional logit: set $S_i = I$
- Mixing both allows for heterogeneous effects of product characteristics, based on observable characteristics
- Or to impose additional structure

The choice probability and estimation on individual data I

- Assumption that translates data on choices to primitives of the model: individuals choose the option with the highest utility
  - $d_{ij} = 1$ if and only if

$$u_{ij} + \epsilon_{ij} > u_{ij'} + \epsilon_{ij'} \text{ for all } j \neq j'$$

- With the extreme value assumption, we derive the following choice probabilities

$$Pr(d_{ij} = 1 | S_i, X_j) = \frac{exp(u_{ij})}{\sum_{j'} exp(u_{ij'})}$$

- with individual data, this is the input we need in the loglikelihood function:

$$lnL = \sum_i \sum_j d_{ij} ln \frac{exp(u_{ij})}{\sum_{j'} exp(u_{ij'})}$$

Normalizations I

- Normalization 1: only differences in utility are identified
  - We can add an arbitrary number to both sides of the inequality
  - In practice, we usually assume the utility of one alternative is 0, then parameters identify differences in utility

Normalizations II

- School choice example with 2 types of schools: $j = 1$ a private school, and $j = 0$ a public school

$$u_{i1} + \epsilon_{i1} = \alpha_1^0 + \alpha_1^{IQ} IQ_i + \epsilon_{i1}$$
$$u_{i0} + \epsilon_{i0} = \alpha_0^0 + \alpha_0^{IQ} IQ_i + \epsilon_{i0}$$

$$Pr(d_{i1} = 1 | S_i, X_j) = \frac{exp(\alpha_1^0 + \alpha_1^{IQ} IQ_i)}{exp(\alpha_0^0 + \alpha_0^{IQ} IQ_i) + exp(\alpha_1^0 + \alpha_1^{IQ} IQ_i)}$$
$$= \frac{exp((\alpha_1^0 - \alpha_0^0) + (\alpha_1^{IQ} - \alpha_0^{IQ}) IQ_i)}{1 + exp((\alpha_1^0 - \alpha_0^0) + (\alpha_1^{IQ} - \alpha_0^{IQ}) IQ_i)}$$

Normalizations III

- As the choice probability only depends on differences in parameters, we might as well assume $\alpha_0 = 0$ and interpret other parameters as the change in utility from choosing a private over a public school
- What about the car example?

- Normalization 2: the scale of utility is irrelevant
  - Also multiplying both sides by something >0 doesn't affect the inequality
  - In practice, one usually normalizes the variance of the error terms to derive the choice probability, we did this by implicitly assuming a variance of $\frac{\pi^2}{6}$
  - When a money-metric variable enters the utility function, it can be useful to normalize the utility function with respect to euros instead and estimate the variance

Interpretation of parameters

- Estimated parameters depend on the scale
- The sign, significance and relative importance can be interpreted
- If one parameter measures the marginal utility of income (e.g. a price coefficient), dividing all parameters by that gives you a willingness to pay measure in euros
- Another way to interpret the results is to calculate average marginal effects
  - Predict choices in status quo but characteristic of interest set to 0
  - Increase a characteristic by one unit
  - Predict choices in new scenario
  - Calculate difference and average over all agents

Counterfactual choices I

- With the estimated model, we can do counterfactuals
- Crucial assumption: characteristics and parameters must be policy-invariant
- Example: the effect of a 10% price decrease of all cars:

$$Pr(d_{ij} = 1|X_j) = \frac{exp(\eta^0 + \eta^{MPG} MPG_j + (1 - 0.10)\eta^P P_j)}{1 + \sum_{j'} exp(\eta^0 + \eta^{MPG} MPG_{j'} + (1 - 0.10)\eta^P P_{j'})}$$

- Is this the same as predicting the impact of a subsidy?
- What if this is part of a budget-neutral government program that taxes the most gas-guzzling cars when they enter highly populated areas?

Counterfactual choices II

- Insights for structural models in general:
  - Several (general) equilibrium effects often assumed away so think about their importance and in which way a potential bias will go
  - A model is never 100% "structural", need to think about what will most likely remain stable in the counterfactual.
  - Good model $\neq$ most detailed model, it's the model that endogenizes what is expected to change in your counterfactual of interest
  - Always judge a model by what you need it for

Consumer surplus I

- Apart from counterfactual choices, we can also assess the change in consumer surplus

$$CS_i = \frac{1}{\alpha} max_{j'}(u_{ij'} + \epsilon_{ij'})$$

- with α the marginal utility of income, e.g. a price coefficient, to scale consumer surplus in euros

- Since we do not observe ϵ, we instead focus on

$$E(CS_i) = \frac{1}{\alpha} E[max_{j'}(u_{ij'} + \epsilon_{ij'})]$$

Consumer surplus II

- for the extreme value type 1 case, it can be shown that the change in consumer surplus is

$$\triangle\ E(CS_i) = \frac{1}{\alpha}(ln \sum_{j'} exp(u^1_{ij'}) - ln \sum_{j'} exp(u^0_{ij'}))$$

- with $u^1_{ij'}$ referring to the counterfactual and $u^0_{ij'}$ to the status quo
- note that this formula requires α to be independent of income

Independence of Irrelevant Alternatives I

- Unobserved heterogeneity enters only through the iid extreme value type 1 on ϵ_{ij}
  - No unobserved heterogeneity in valuation of product characteristics
  - No correlation of unobserved tastes over different alternatives
- This implies Independence of Irrelevant Alternatives (IIA)

$$\frac{Pr(d_{ij} = 1|X_j, S_i)}{Pr(d_{ij'} = 1|X_j, S_i)} = exp(u_{ij} - u_{ij'})$$

- The ratio of two probabilities does not depend on a third alternative

Independence of Irrelevant Alternatives II

- Red-bus-blue-bus problem
  - Traveler chooses between blue bus or car and we know that they are divided 50-50-> ratio=1
  - Suppose a red bus is introduced and travelers don't care about color, then the probability of both buses should be the same
  - Because of IIA the ratio blue bus - car remains (also) 1
  - Therefore all probabilities must be 1/3, while we would expect 50% bus (25% blue, 25% red) and 50% car
- This shows that logit models are unfit to study substitution patterns, this is often crucial for counterfactuals
  - Take for example an increase in prices for a Ferrari, you want the model to predict an increase in demand for Porsche but not for Dacia

Nested logit I

- One way to reduce this problem is to add demographics (S_i) (why?)
- But this is not always available or sufficient so instead we could assume a different distribution of unobservables in the model
- This often leads to models without closed form solutions for the choice probabilities
- However, the nested logit still allows for this, while being able to solve the red-bus-blue-bus problem
- A researcher first has to declare a nest (e.g. "bus" or "fancy cars"), within each nest ϵ_{ij} is allowed to be correlated

Nested logit II

- Utility is now (notation from Berry (1994))

$$u_{ij} + \zeta_{ik} + (1 - \sigma)\epsilon_{ij}$$

- with $\sigma \in [0, 1]$ a parameter that approximates the correlation within nests (distribution ζ_{ik} depends on σ)
  - $\sigma = 0$: back to standard logit
  - $\sigma = 1$: products within nests are perfect substitutes

Nested logit III

- Choice probability still has closed form solution

$$Pr(d_{ij} = 1|X_j, S_i) = \frac{exp(u_{ij}/(1-\sigma))(\sum_{j'\in B_k} exp(u_{ij'}/(1-\sigma)))^{-\sigma}}{\sum_{k'}(\sum_{j'\in B_{k'}} exp(u_{ij'}/(1-\sigma)))^{(1-\sigma)}}$$

- σ can be different for each nest
- Several extensions are possible: multi-level nests, overlapping nests,...
- These are members of the Generalized Extreme Value (GEV) family
- Despite increased computer power, obtaining tractable, yet flexible models remains an important area of research (see e.g. Grigolon (2020) or Fosgerau et al. (2021))

Random coefficients logit I

- It is often difficult (and arbitrary) to define a nest
  - (see ongoing work Almagro & Manresa for data-driven way)
- If one is willing to accept a model without closed form solution for the choice probabilities, the model can be made even more flexible
- A popular alternative is to allow for unobserved heterogeneity in the "taste" parameters η
- Take the model without observed consumer heterogeneity (S_i)

$$u_{ij} + \epsilon_{ij} = \eta^0 + \eta^{MPG} MPG_j + \eta^P P_j + \epsilon_{ij}$$

- and generalize to

$$u_{ij} + \epsilon_{ij} = \eta_i^0 + \eta_i^{MPG} MPG_j + \eta_i^P P_j + \epsilon_{ij}$$

Random coefficients logit II

- with $\eta_i = (\eta_{ij}^0, \eta_i^{MPG}, \eta_i^P)$ following a (log)normal distribution with estimated means and covariance matrix
- As with observed heterogeneity, this creates flexible subsitution patterns
  - E.g. someone with η_i^P close to 0 is likely to switch between Ferrari and Porsche, while someone with a strongly negative η_i^P will switch between Dacia and Kia
- Can easily add interactions with S_i to also allow for observed heterogeneity

Random coefficients logit III

- Choice probabilities have to be simulated which increases computational complexity and estimation time
  - See website of Kenneth Train for matlab/R/Gauss code, or use user-written STATA command "mixlogit"
- For the same reason, restrictions are often imposed on the covariance matrix
- Can also assume finite mixture of types with each type having its own η_i (more popular with dynamics, will come back to this)
- An alternative for the RC logit model that also requires simulation is to assume a normal distribution on ϵ_{ij}, with estimated covariances between alternatives (ch. 5 in Train (2009))

The use of aggregate data I

- Similar models have been proposed to estimate demand using market share data
- This is very popular in IO where people are interested in estimating (cross-)price elasticities between a large number of differentiated goods
- Start again from the conditional logit, but add an "unobserved product characteristic", (or demand shock, or quality,..) ξ_j

$$u_{ij} + \epsilon_{ij} = u_j + \epsilon_{ij} = X_j'\eta + \xi_j + \epsilon_{ij}$$

- $X_j'\eta + \xi_j$ is also known as the "mean utility", often written as δ_j
- Berry (1994) shows that this model can be estimated directly using a linear regression on market share data

The use of aggregate data II

$$Pr(d_{ij} = 1|\delta) = \frac{exp(u_j)}{\sum_{j'} exp(u_{j'})}$$

(normalization of outside good)

$$Pr(d_{i0} = 1|\delta) = \frac{exp(u_0)}{\sum_{j'} exp(u_{j'})} = \frac{1}{\sum_{j'} exp(u_{j'})}$$

then

$$ln(Pr(d_{ij} = 1|\delta)) - ln(Pr(d_{i0} = 1|\delta)) = u_j = X_j'\eta + \xi_j$$

Because of the inclusion of ξ_j, we can fit market shares exactly

$$ln(s_j) - ln(s_0) = X_j'\eta + \xi_j$$

The use of aggregate data III

- this is a simple OLS regression with dependent variable $ln(s_j) - ln(s_0)$, covariates X_j and error term ξ_j
- In contrast to the extreme value error ϵ_{ij}, the unobserved product characteristic ξ_j is easy to allow for heteroskedasticity and arbitrary correlation (use robust or clustered standard errors), and, most importantly, to account for endogeneity by using 2SLS
  - Main example: prices in X are correlated with unobserved quality, so we need a cost-shifter as instrument
- As with individual data, this model is easy to extend to a nested logit and (less easy) to a random coefficients logit, the latter is also known as the BLP (1995) model
  - See Nevo (2000) for a good "practitioner's guide" for BLP

The use of aggregate data IV

- Note that the number of observations is equal to the number of products, not the number of agents. To gain power, it is usually estimated over multiple markets or time periods (t):

$$ln(s_{jt}) - ln(s_{0t}) = X_{jt}'\eta + \xi_{jt}$$

- This model can also be used with individual data by estimating fixed effects δ_j using maximum likelihood in a first stage and regressing this on product characteristics in a second stage, this way endogeneity problems can be solved while allowing for individual heterogeneity

The use of aggregate data V

1. Likelihood function first stage

$$lnL = \sum_i \sum_j d_{ij} ln \frac{exp(\delta_j + S_i' \omega X_j)}{1 + \sum_{j' \neq 0} exp(\delta_{j'} + S_i' \omega X_j)}$$

2. Linear regression second stage

$$\delta_j = X_j' \eta + \xi_j$$

3. Correct standard errors for estimation error or estimate jointly using GMM

- Or can use a control function approach (Petrin and Train, 2010)

References I

Jaap H. Abbring and Øystein Daljord.
Identifying the discount factor in dynamic discrete choice models.
Quantitative Economics, 11(2):471–501, 2020.

Victor Aguirregabiria and Pedro Mira.
Swapping the nested fixed point algorithm: A class of estimators for discrete Markov decision models.
Econometrica, 70(4):1519–1543, 2002.

Victor Aguirregabiria and Pedro Mira.
Dynamic discrete choice structural models: A survey.
Journal of Econometrics, 156(1):38–67, May 2010.

References II

- Milena Almagro and Tomas Domınguez-Iino.
  Location Sorting and Endogenous Amenities: Evidence from Amsterdam.
  Working paper, 2021.

- Peter Arcidiacono, Esteban Aucejo, Arnaud Maurel, and Tyler Ransom.
  College Attrition and the Dynamics of Information Revelation.
  Working paper, 2016.

- Peter Arcidiacono and Paul B. Ellickson.
  Practical Methods for Estimation of Dynamic Discrete Choice Models.
  Annual Review of Economics, 3:363–394, 2011.

References III

- Peter Arcidiacono and Robert A. Miller.
  Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity.
  Econometrica, 79(6):1823–1867, 2011.

- Jared Ashworth, V Joseph Hotz, Arnaud Maurel, and Tyler Ransom.
  Changes across Cohorts in Wage Returns to Schooling and Early Work Experiences.
  Journal of Labor Economics, (forthcoming), 2021.

- Christian Belzil.
  The return to schooling in structural dynamic models: a survey.
  European Economic Review, 51(5):1059–1105, July 2007.

References IV

Steven Berry.
Estimating Discrete-Choice Models of Product Differentiation.
The RAND Journal of Economics, 25(2):242–262, 1994.

Steven Berry, James Levinsohn, and Ariel Pakes.
Automobile Prices in Market Equilibrium.
Econometrica, 63(4):841, July 1995.

Olivier De Groote.
A Dynamic Model of Effort Choice in High School.
TSE working paper - 1002, 2020.

Koen Declercq and Frank Verboven.
Enrollment and degree completion in higher education without admission standards.
Economics of Education Review, 66:223–244, October 2018.

References V

Mogens Fosgerau, Julien Monardo, and André de Palma.
The Inverse Product Differentiation Logit Model.
Working paper, January 2021.

Laura Grigolon.
Blurred boundaries: a flexible approach for segmentation applied to the car market.
Quantitative Economics, (forthcoming), 2020.

V. Joseph Hotz and Robert A. Miller.
Conditional choice probabilities and the estimation of dynamic models.
The Review of Economic Studies, 60(3):497–529, 1993.

References VI

Yingyao Hu and Matthew Shum.
Nonparametric identification of dynamic models with unobserved state variables.
Journal of Econometrics, 171(1):32–44, November 2012.

Myrto Kalouptsidi, Paul T. Scott, and Eduardo Souza-Rodrigues.
Linear IV regression estimators for structural dynamic discrete choice models.
Journal of Econometrics, 222(1):778–804, May 2021.

Michael P. Keane.
Structural vs. atheoretic approaches to econometrics.
Journal of Econometrics, 156(1):3–20, May 2010.

References VII

Michael P. Keane, Petra E. Todd, and Kenneth I. Wolpin.
The Structural Estimation of Behavioral Models: Discrete Choice Dynamic Programming Methods and Applications.
In *Handbook of Labor Economics*, volume 4, pages 331–461. Elsevier, 2011.

Michael P. Keane and Kenneth I. Wolpin.
The Career Decisions of Young Men.
Journal of Political Economy, 105(3):473–522, June 1997.

Thierry Magnac and David Thesmar.
Identifying dynamic discrete decision processes.
Econometrica, 70(2):801–816, March 2002.

References VIII

Alvin Murphy.
A Dynamic Model of Housing Supply.
American Economic Journal: Economic Policy, 10(4):243–267, November 2018.

Aviv Nevo.
A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand.
Journal of Economics & Management Strategy, 9(4):513–548, 2000.

Amil Petrin and Kenneth Train.
A control function approach to endogeneity in consumer choice models.
Journal of marketing research, 47(1):3–13, 2010.

References IX

Lionel Robins.
Essay on the Nature and Significance of Economic Science.
Macmillan, London, 1932.

John Rust.
Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher.
Econometrica, 55(5):999–1033, September 1987.

John Rust.
Comments on: "Structural vs. atheoretic approaches to econometrics" by Michael Keane.
Journal of Econometrics, 156(1):21–24, May 2010.

References X

Paul T. Scott.
Dynamic Discrete Choice Estimation of Agricultural Land Use.
Working paper, 2013.

Sharon Traiberman.
Occupations and Import Competition: Evidence from Denmark.
American Economic Review, 109(12):4260–4301, December 2019.

Kenneth Train.
Discrete choice methods with simulation.
Cambridge University Press, Cambridge; New York, 2009.