Dynamic discrete choice models: CCP methods

Olivier De Groote

Toulouse School of Economics - University of Toulouse Capitole

Penn guest lecture, April 2024

Outline

CCP estimation

- One-step-ahead
- Finite dependence
- Expectational errors

Outline

Avoid solving the model

- The computational complexity of dynamic models quickly increases
- Need to find parameters and solve an economic model at the same time
- If solving the model takes a long time, it prevents us from enriching the model to make it more realistic
- The Rust (1987) framework was used to develop estimation methods that do not require solving the model: CCP methods (introduced by Hotz and Miller (1993))
- I highly recommend (and closely follow) the review: Arcidiacono, P., and Ellickson, P. 2011. "Practical Methods for Estimation of Dynamic Discrete Choice Models". *Annual Review of Economics* 3: 363–94.

CCP estimation with one-period-ahead CCPs I

- Conditional value function

$$v_{ijt} = u_j(x_{it}) + \beta \int \bar{V}_{t+1}(x_{it+1}) f(x_{it+1}|x_{it}, d_{it} = j) dx_{it+1}$$

- Slightly different notation than Arcidiacono and Ellickson
  - I subscript by alternative j instead of v and u depending explicitly on the choice d
  - I add the i subscript
- Static agents choose the option with the highest $u_j(x_{it}) + \epsilon_{ijt}$ -> easy to put in likelihood function
- Dynamic agents choose the option with the highest value of $v_{ijt} + \epsilon_{ijt}$ -> as easy, once we get rid of $+\beta$...

CCP estimation with one-period-ahead CCPs II

- With ϵ_{ijt} extreme value type 1, we know the logsum expression for the ex ante value function

$$\bar{V}_{t+1}(x_{it+1}) = \gamma + ln \sum_{j'} exp(v_{ij't+1})$$

- At the same time, we also have the standard logit probabilities

$$Pr(d_{it+1} = j^*|x_{it+1}) = \frac{exp(v_{ij^*t+1})}{\sum_{j'} exp(v_{ij't+1})}$$

- Taking logs

$$lnPr(d_{it+1} = j^*|x_{it+1}) = v_{ij^*t+1} - ln \sum_{j'} exp(v_{ij't+1})$$

CCP estimation with one-period-ahead CCPs III

- Rearrange

$$ln \sum_{j'} exp(v_{ij't+1}) = v_{ij^*t+1} - lnPr(d_{it+1} = j^*|x_{it+1})$$

- Substituting this in the ex ante value function

$$\bar{V}_{t+1}(x_{it+1}) = \gamma + v_{ij^*t+1} - lnPr(d_{it+1} = j^*|x_{it+1})$$

- The expected value of behaving optimally in $t+1$ depends only on the value of an arbitrary choice j^*, and a nonnegative correction term $-lnPr(d_{it+1} = j^*|x_{it+1})$
- $Pr(d_{it+1} = j^*|x_{it+1})$ is data (or can at least be predicted)
- So instead of solving the model, we only need to know the conditional value function of one arbitrary option j^*

CCP estimation with one-period-ahead CCPs IV

- In general, this remains complicated as this will again depend on $\bar{V}_{t+2}$ etc...
- But take for example a terminal action (e.g. a retirement decision, buying a durable good,..). In this case we can write

$$v_{ij^*t+1} = u_{ij^*t+1}$$

- (or, depending on interpretation, $\frac{u_{ij^*t+1}}{1-\beta}$)
- This can be parameterized as a function of state variable x_{it+1}, or normalized to 0.
- In the latter case we get

$$v_{ijt} = u_j(x_{it}) + \beta \int (\gamma - lnPr(d_{it+1} = j^* | x_{it+1})) f(x_{it+1} | x_{it}, d_{it} = j) dx_{it+1}$$

CCP estimation with one-period-ahead CCPs V

- As in Rust (1987), we get state transitions $f(x_{it+1}|x_{it}, d_{it} = j)$ in a first stage, now we just need to get $Pr(d_{it+1} = j^*|x_{it+1})$ ("CCPs") in a first stage too

- The model becomes a static model with a prespecified correction term: no need to solve anything!

CCP estimation with one-period-ahead CCPs VI

- Another example where CCPs are useful is when there is a choice that "resets" the state

- To see this, first note that the logit probabilities can be written using differenced value functions

$$Pr(d_{it} = j|x_{it}) = \frac{exp(v_{ijt})}{\sum_{j'} exp(v_{ij't})} = \frac{exp(v_{ijt} - v_{i0t})}{1 + \sum_{j' \backslash 0} exp(v_{ij't} - v_{i0t})}$$

CCP estimation with one-period-ahead CCPs VII

- Without a terminal action, every v_{ijt} will be of the form

$$v_{ijt} = u_j(x_{it}) + \beta \int (\gamma + v_{ij^*t+1} - lnPr(d_{it+1} = j^* | x_{it+1})) f(x_{it+1} | x_{it}, d_{it} = j) dx_{it+1}$$

- with

$$v_{ij^*t+1} = u_{j^*}(x_{it+1}) + \beta \int \bar{V}_{t+2}(x_{it+2}) f(x_{it+2} | x_{it+1}, d_{it+1} = j^*) dx_{it+2}$$

CCP estimation with one-period-ahead CCPs VIII

- Now assume j^*at time $t+1$ is a "renewal" action, meaning it resets the state such that

$$\int f(x_{it+1}|x_{it}, d_{it}=j) f(x_{it+2}|x_{it+1}, d_{it+1}=j^*) dx_{it+1}$$
$$= \int f(x_{it+1}|x_{it}, d_{it}=j') f(x_{it+2}|x_{it+1}, d_{it+1}=j^*) dx_{it+1}$$

- in words: the state at $t+2$ will not depend on the choice made today, if tomorrow we choose a renewal action
  - Example: in the Rust bus engine model, replacement in $t+1$ is a renewal action because it sets mileage to 0, regardless of what Harold Zurcher chooses to do today

CCP estimation with one-period-ahead CCPs IX

- Following this assumption $\beta \int \bar{V}_{t+2}(x_{it+2}) f(x_{it+2}|x_{it+1}, d_{it+1} = j^*) dx_{it+2}$ drops out of $v_{ijt} - v_{i0t}$ and we can write the logit probabilities without a future value term

$$
\begin{aligned}
v_{ijt} - v_{i0t} = {} & u_j(x_{it}) + \beta \int (\gamma + u_{ij*\,t+1}(x_{it+1}) - lnPr(d_{it+1} = j^*|x_{it+1})) f(x_{it+1}|x_{it}, d_{it} = j) dx_{it+1} \\
& - u_0(x_{it}) - \beta \int (\gamma + u_{ij*\,t+1}(x_{it+1}) - lnPr(d_{it+1} = j^*|x_{it+1})) f(x_{it+1}|x_{it}, d_{it} = 0) dx_{it+1}
\end{aligned}
$$

Outline

CCP estimation with more CCPs I

- If resetting the state is not possible by an action in the next period, it might work after 2,3,.. periods
- An example is Arcidiacono et al. (2016)
  - Choose between several work/college options or staying home
  - Primitives depend on accumulated work experience and past choice
  - Impossible to remove $\bar{V}_{t+2}(x_{it+2})$ in $v_{ijt} - v_{i0t}$
- Solution: solve one period further
- Consider 2 sequences of choices:
  - Choose j now, 0 in $t+1$ and 0 in $t+2$
  - Choose 0 now, j in $t+1$ and 0 in $t+2$
- In $t+3$ both have the same past choice and the same experience -> $\bar{V}_{t+3}(x_{it+3})$ is the same

CCP estimation with more CCPs II

- Need to rewrite v_{ijt} and v_{i0t} such that they both end with a term $\bar{V}_{t+3}(x_{it+3})$

- Use $\bar{V}_{t+2}(x_{it+2}) = \gamma + v_{ij^*t+2} - lnPr(d_{it+2} = j^*|x_{it+2})$ with $v_{ij^*t+2} = u_{j^*}(x_{it+2}) + \beta \int \bar{V}_{t+3}(x_{it+3}) f(x_{it+3}|x_{it+2}, d_{it+2} = j^*) dx_{it+3}$

(notation is not great here as the arbitrary choice j^* differs in each sequence and time period)

CCP estimation without finite dependence

- CCP estimation can still be useful in one of the following ways
  - In stationary problems, the $\bar{V}()$ can be written as a function of CCPs and utility parameters by inverting a matrix instead of looking for a fixed point as in Rust (1987)
  - We can write the problem using CCPs until the end (or sufficiently far into the future if infinite horizon)
- CCP estimators are less efficient, one can update the CCPs using the model to regain efficiency (Aguirregabiria and Mira (2002))
- Standard errors are not correct when CCPs have to be estimated first

Counterfactual?

- CCPs not valid in counterfactuals
- But only need to solve those once...
- Can also choose smart counterfactuals
  - Arcidiacono et al. (2016): no learning about ability
  - De Groote & Rho: truth-revealing mechanisms
  - De Groote & Verboven (2019): indifferent agents

Outline

CCP estimation

- One-step-ahead
- Finite dependence
- Expectational errors

CCP estimation to relax assumptions I

- CCP estimation was also extended to relax assumption on agents' expectations in Scott (2013)
- In De Groote & Verboven (2019), we estimate such a model using market shares as CCPs
- Model the decision of households when and how much to invest in a solar panel
- Main question: what would be the change in government expenditures if subsidization policy mainly reduced investment cost instead of increasing future benefits?

Model (linear version of the paper) I

- Each month t, every household that has not yet adopted a PV decides to either:
  - Don't adopt ($j = 0$)
  - Adopt 2kw, 4kw, 6kw, 8kw or 10kw: ($j = 1, ..., 5$)
- Conditional value function of adoption (net of ε_{ijt})

$$v_{i,j,t} = \delta_{j,t} = x_{j,t}\gamma - \alpha p_{jt} + \alpha\theta b_{jt} + \xi_{jt}$$

- x are observed characteristics (here just choice-specific constants)
- ξ_{jt} is a demand shock
- p_{jt} is the investment cost
- b_{jt} is the present value of investment benefits
- (Paper also adds micro-level heterogeneity in an extension)

Model (linear version of the paper) II

- Conditional value of not adopting ($j = 0$):

$$\delta_{0,t} = u_{0,t} + \beta E_t \overline{V}_{t+1}$$

- with $\overline{V}_{t+1} = 0.577 + \ln \sum_{j=0}^{J} \exp\left(\delta_{j,t+1}\right)$.
- We follow important insight of Scott (2013), instead of specifying households expectations of the future $\left(E_t \overline{V}_{t+1}\right)$, we define a prediction error

$$\eta_t \equiv \overline{V}_{t+1} - E_t \overline{V}_{t+1}$$

and use realizations of the future instead. In estimation, we can treat η_t in a similar way as the demand shocks ξ_{jt}, we don't need to know them, we simply need instruments that are not correlated with them.

Model (linear version of the paper) III

- Normalize $u_{0,t} + 0.577\beta = 0$ such that

$$\delta_{0,t} = \beta \left(\ln \sum_{j=0}^{J} \exp\left(\delta_{j,t+1}\right) - \eta_t \right).$$

- Now we apply CCP using $j = 1$ as a terminal action
- Because we use realized value functions and not expected value functions, the CCP is simply the market share:

$$S_{1,t+1} = \frac{\exp\left(\delta_{1,t+1}\right)}{\sum_{j=0}^{J} \exp\left(\delta_{j,t+1}\right)}$$

$$ln \sum_{j=0}^{J} \exp\left(\delta_{j,t+1}\right) = \delta_{1,t+1} - lnS_{1,t+1}$$

Model (linear version of the paper) IV

- Substituting back into the mean utility of not adopting:

$$\delta_{0,t} = \beta \left(\delta_{1,t+1} - \ln S_{1,t+1} - \eta_t \right)$$

- Now we do the same as in Berry (1994), it still holds that

$$\ln S_{j,t} / S_{0,t} = \delta_{j,t} - \delta_{0,t}$$

- Because

$$S_{j,t} = \frac{\exp\left(\delta_{j,t}\right)}{\sum_{j'=0}^{J} \exp\left(\delta_{j',t}\right)}$$

and

$$S_{0,t} = \frac{\exp\left(\delta_{0,t}\right)}{\sum_{j'=0}^{J} \exp\left(\delta_{j',t}\right)}$$

Model (linear version of the paper) V

- But the "mean utilities" look different now

  $$\ln S_{j,t}/S_{0,t} = (x_{j,t} - \beta x_{1,t+1})\gamma - \alpha(p_{j,t} - \beta p_{1,t+1}) + \alpha\theta(b_{j,t} - \beta b_{1,t+1}) + \beta \ln S_{1,t+1} + e_{j,t}$$

  where $e_{j,t} \equiv \xi_{j,t} - \beta(\xi_{1,t+1} - \eta_t)$

- Fix β and estimate using 2SLS with instruments
  - product characteristics to identify γ
  - price of Chinese PV modules to identify α
  - present value of benefits to identify θ
- In the paper, we actually estimate β, and no θ, by proposing a non-linear version of this model

Budgettary implication

- Households perceived the investment benefits as follows

$$\frac{1-(0.9884)^R}{1-0.9884}GCC_{j,t}$$

- But the government is paying

$$\frac{1-(0.9975)^R}{1-0.9975}GCC_{j,t}$$

- Adding some further technical corrections, we find that an upfront subsidy would have implied a saving of 51% or €1.9 billion (€700/hh)

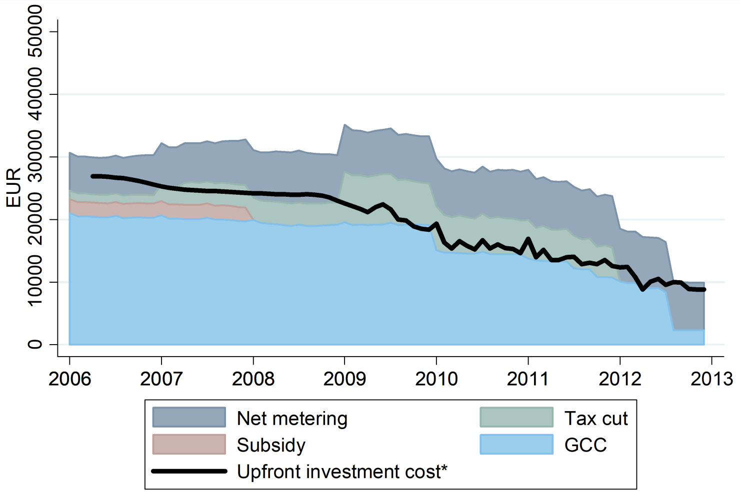
Other examples of ECCP estimators

- Because the regression to estimate can be interpreted as a discretized version of an Euler equation, the method introduced by Scott (2013) is now often called ECCP estimators, see Kalouptsidi et al. (2021) for general framework
- Also works with individual data: calculate choice probabilities for types of individuals instead of market shares
- See Traiberman (2019) for application in labor/trade
- See Almagro & Domınguez-Iino (2021) for application in housing

Extra: Costs and benefits of a PV (2006-2012) I

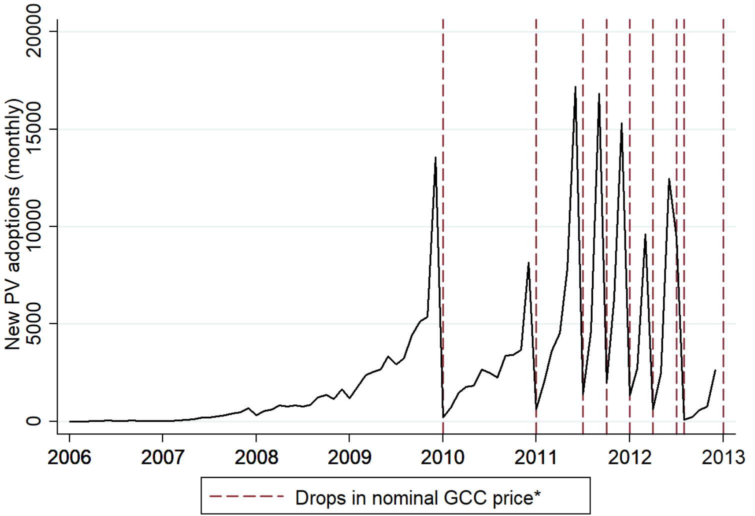
- Upfront investment subsidies and tax cuts
- Future benefits from net metering
  - Electricity bill $=$ electricity price $\times$ (consumption-production)
- Future production subsidies: Green Current Certificates (GCCs)
  - Households get a GCC per MWh for fixed number of years (mostly 20)
  - GCCs can be sold to grid operator at a guaranteed price
  - The price guarantee is fixed at the moment of the investment

Extra: Costs and benefits of a PV of 4 kW (2006-2012)

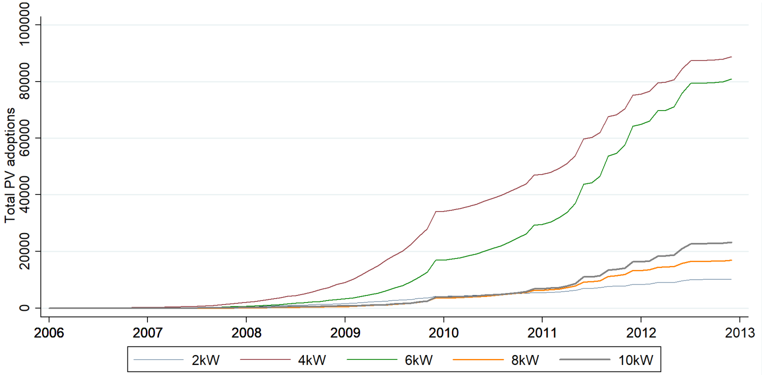


Upfront investment cost*: predicted values using German price index EUPD before May 2009
Real interest rate used to calculate present values = 3%

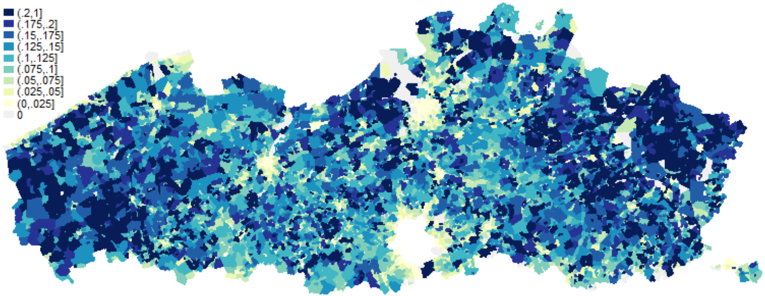
Extra: Evolution of new PV adoption and drops in GCC price



Extra: Evolution of total PV adoption by capacity



Extra: PV adoption across Flanders



Extra: Local market heterogeneity I

- Heterogeneity only entered through iid ε_{ijt}
- In static model we know this is problematic because of correlation between similar goods
- Similar solutions exist in a dynamic model, CCP still holds for extension like nested logit
- But in dynamic models, also nested logit doesn't solve another issue: correlation over time!
- Possible solutions in the literature: random coefficients as in BLP (see e.g. Gowrisankaran and Rysman, 2012) or types (see e.g. Scott, 2013)
- Alternative in this paper: test robustness by incorporating local market heterogeneity using micro-moments

Extra: Local market heterogeneity II

- Add covariates of 9,182 local markets m with an average of 295 households
- Add local market-specific term to conditional value of adoption

$$v_{i,j,t} = \delta_{j,t} + \mu_{m,j,t}$$

$$= \delta_{j,t} + w_{j,t}\lambda_m$$

where $w_{j,t}\lambda_m = w_{j,t}\Lambda D_m$ and Λ contains interactions between product characteristics $w_{j,t}$ and local market variables D_m

- We allow for a large set of local market fixed effects to control for unobserved heterogeneity
- And interactions of demographics with capacity and price

Extra: Local market heterogeneity III

- Estimation of Λ by adding micro-moments to GMM estimator
  - Micro-moments match the observed covariances between the demographic variables and product characteristics to the models predictions, as well as the total number of adopters in each market at the end of the sample
  - They can be derived from a likelihood model that maximizes the likelihood to observe the local market shares

References I

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.

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

Peter Arcidiacono and Paul B. Ellickson.
Practical methods for estimation of dynamic discrete choice models.
Annual Review of Economics, 3(1):363–394, March 2011.

References II

Olivier De Groote and Frank Verboven.
Subsidies and Time Discounting in New Technology Adoption: Evidence from Solar Photovoltaic Systems.
American Economic Review, 109(6):2137–2172, June 2019.

Gautam Gowrisankaran and Marc Rysman.
Dynamics of Consumer Demand for New Durable Goods.
Journal of Political Economy, 120(6):1173–1219, December 2012.

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 III

- 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.
- John Rust.
  Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher.
  Econometrica, 55(5):999–1033, September 1987.
- Paul T. Scott.
  Dynamic Discrete Choice Estimation of Agricultural Land Use.
  Working paper, 2013.

References IV

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