A dynamic discrete choice model of land use conversion

May 4, 2021

1 Theory

1.1 The decision problem

This section adapts the formulation of the dynamic discrete choice problem described in Kalouptsidi et al. (2020) to a dynamic discrete choice over land use and land use change. In each period, the landowner chooses whether to keep their land in its current land use or to convert it to a different use, where this decision is made to maximize expected profits over time.

Profits in each period depend on the state $\mathbf{s} = \{k_{it}, w_{mt}, \epsilon_{it}, \eta_{mt}\}$, where variables indexed by i vary at the individual parcel level and variables indexed only by m vary at the market-level. We define market m to consist of all parcels of quality q within a given county c. The variables (k_{it}, w_{mt}) are observed by the econometrician, whereas $(\epsilon_{it}, \eta_{mt})$ are not. The variable k_{it} denotes the current land use at the beginning of period t. While the evolution of x_{it} is controlled by the landowner, w_{mt} contains variables that are observed at the market level and are thus unaffected by landowner choices (assuming the market is sufficiently large); for example, w_{mt} contains input and output prices for commodities produced on land in each use. Observed and unobserved market-level state variables are collected into the vector $\omega_{mt} = (w_{mt}, \eta_{mt})$.

Landowners' profits in each period depend on parcel and market-level state variables, and the choice of land use j_{it} in period t, and are written:

$$\Pi(j_{it}, \mathbf{s}_{imt}) = \bar{\pi}_m(j_{it}, k_{it}, w_{mt}) + \xi(j_{it}, k_{it}, \omega_{mt}) + \epsilon_{jimt}. \tag{1}$$

Therefore, landowner profits are the sum of $\bar{\pi}$, a function of observed variables, ξ, a function that depends also on unobserved market-level variables, and the disturbance term ϵ_{jimt}, which we assume follows a Type I extreme value distribution. We define π such that $\pi \equiv \bar{\pi} + \xi$, and $\Pi = \pi + \epsilon$.

Consistent with our data and previous literature, we assume that urban use is an absorbing state; owners of land in non-urban uses who develop their land receive a sales price equal to the net present value of urban rents from land perpetuity. Note: This may not be necessary. We denote urban use as J and specify the observable portion of landowner profits as:

$$\bar{\pi}_m(j_{it}, x_{it}, w_{mt}) = \begin{cases} R_m(j_{it}, k_{it}, w_{mt}) - C_{mkj} & j \neq J \\ \frac{R_m(j_{it}, k_i it, w_{mt})}{1-r} - C_{mkj} & j = J \end{cases} \tag{2}$$

where C_{kJ} represents the cost of converting land into urban use, and r is the interest rate.

For land uses $j \neq J$, we follow Lubowski et al. (2006) in specifying $R_j(w_{mt}, q_i)$ according to:

$$R_m(j_{it}, k_{it}, w_{mt}) = \theta^0_{jk} P_{jct} + \sum_{\bar{q}=2}^{6} \theta^{\bar{q}}_{jk} (\mathbb{1}(q_m = \bar{q}) \times P_{jct}) \tag{3}$$

where P_{jct} represents the measured return to land use j within county c at time t, and q_m represents the land condition class, a measure of land quality, of parcels in market m. Note: Lubowski et al. (2006) include separate values of θ_1 and θ_2 for each starting land use k. Why? Also, why are the values of θ subscripted by t? For now, following from this, I've subscripted the R and C functions by t, but I'm not totally sure why this makes sense. Again following Lubowski et al. (2006), we specify conversion costs as:

$$C_{kjt}(q_i) = \eta^0_{kjt} + \sum_{\bar{q}=2}^{6} \eta^{\bar{q}}_{jk} \mathbb{1}(q_m = \bar{q}) \tag{4}$$

where conversion costs are allowed to vary with parcel quality.

In each period the landowner chooses whether to convert their land into an alternative use based on the profits they would receive from doing so, and the expected implications for profits in future periods. This decision problem can be described recursively using the Bellman equation:

$$V(\mathbf{s}_{imt}) = \max_{j_{it}} \pi(j_{it}, k_{it}, \omega_{mt}) + \epsilon_{jimt} + \beta \mathrm{E}\Big[V_{t+1}(\mathbf{s}_{imt+1}) | j_{it}, \mathbf{s}_{imt}\Big] \tag{5}$$

Because V is a recursive function, estimating its structural parameters is not straightforward. Arriving at the estimating equation requires first writing V in terms of known functions. The next section defines functions that will be useful in doing so.

1.2 Value functions, choice probabilities, and expectational errors

First, we define the *ex ante* value function $\bar{V}_m(k_{it}, \omega_{mt})$ as the expected utility in the current period, prior to the realization of ϵ:

$$\bar{V}_m(k_{it}, \omega_{mt}) \equiv \mathrm{E}_\epsilon\big[V(k_{it}, \omega_{mt}, \epsilon_{it})\big] = \int V(k_{it}, \omega_{it}, \epsilon_{it}) g(\epsilon_{it}) d\epsilon \tag{6}$$

Next, we define the *conditional* value function as the present discounted value (less ϵ) of choosing j given the current state of parcel i, k_{it}, and market-level state variables ω_{mt}:

$$v_m(j_{it}, k_{it}, \omega_{mt}) = \pi_m(j_{it}, k_{it}, \omega_{mt}) + \beta \mathrm{E}\big[\bar{V}_{mt+1}(k_{it+1}(j_{it}), \omega_{mt+1}) | \omega_{mt})\big] \tag{7}$$

Using the conditional value function, the probability of choosing land use j'—the *conditional choice probability* can be written:

$$p_m(j_{it}|k_{it}, \omega_{mt}) = \int \mathbb{1}\Big\{ \arg\max_{j'}(v(j', k_{it}, \omega_{mt}) + \epsilon(j')) = j_{it} \Big\} g(\epsilon_{it}) d\epsilon \tag{8}$$

From Arcidiacono and Miller (2011) and Hotz and Miller (1993), if ϵ follows a Type I extreme value distribution, then for any arbitrary choice j the ex ante value function can be written as a function of the conditional value function:[1]

$$\bar{V}_m(k_{it}, \omega_{mt}) = v_m(j_{it}, k_{it}, \omega_{mt}) + \gamma - \ln\big(p_m(j_{it}|k_{it}, \omega_{mt})\big). \tag{9}$$

Note that when $p_m(j_{it}|k_{it}, \omega_{mt}) = 1$, the adjustment term equals zero. Therefore, this equation can be interpreted as saying that the ex ante utility of being in state (k_{it}, ω_{mt}) can be written as the conditional value of choosing arbitrary land use j and a term that adjusts for the fact that j may not the optimal choice when in land use j $\big(\gamma - \ln\big(p_m(j_{it}|x_{it}, \omega_{mt})\big)$ (Arcidiacono and Ellickson, 2011). This relationship will be useful for eliminating $\bar{V}$ from the estimating equation.

Finally, we follow **?** and Kalouptsidi et al. (2020) in defining "expectational errors." Expectational errors will be useful in disposing of expectations over $\bar{V}$, allowing us to write $\mathrm{E}\bar{V}$ as the sum of V and its expectational errors. Kalouptsidi et al. (2020) define expectational errors for any function $h(z, \omega)$ and any realization ω^* as:

$$e^h(z', \omega, \omega^*) \equiv \mathrm{E}_{\omega'|\omega}\big[h(z', \omega')|\omega\big] - h(z', \omega^*) \tag{10}$$

Therefore, e^h describes the difference between the expected and realized values $h(\cdot)$ due to realization of ω'. Further, given some action a, they define:

$$e^h(a, z, \omega, \omega^*) \equiv \sum_{x'} e^h(z', \omega, \omega^*) F(z'|a, x, w), \tag{11}$$

the mean expectational error over the agent-determined state variable z' given action a, initial states z and ω, and a realization ω^*.

1.3 ECCP Equations

Using the tools from the last section, we can derive an equation to estimate the structural parameters of landowners' dynamic land use decision. First, we can difference equation 7 across two

[1] Would be worth understanding the derivation of this better.

arbitarily chosen land uses j and a:

$$\begin{aligned} v_m(j_{it}, k_{it}, \omega_{mt}) - v_m(a_{it}, k_{it}, \omega_{mt}) = & \ \pi_m(j_{it}, k_{it}, \omega_{mt}) - \pi_m(a_{it}, k_{it}, \omega_{mt})\big) \\ & + \beta\Big\{\mathrm{E}\big[\bar{V}_{mt+1}(k_{it+1}(j_{it}), \omega_{mt+1})|\omega_{mt}\big] \\ & - \mathrm{E}\big[\bar{V}_{mt+1}(k_{it+1}(a_{it}), \omega_{mt+1})|\omega_{mt}\big]\Big\} \end{aligned}$$

Using the Hotz and Miller (1993) inversion, which writes differences in conditional value functions in terms of conditional choice probabilities, we can rewrite the left-hand side of the equation as:

$$\begin{aligned} \ln\left(\frac{p_m(j_{it}|k_{it}, \omega_{mt})}{p_m(a_{it}|k_{it}, \omega_{mt})}\right) = & \ \big(\pi(j_{it}, x_{it}, w_{mt}) - \pi(a_{it}, x_{it}, w_{mt})\big) \\ & + \beta\Big\{\mathrm{E}\big[\bar{V}_{t+1}(x_{it+1}, \omega_{mt+1})|j_{it}, x_{it}, \omega_{mt})\big] \\ & - \mathrm{E}\big[\bar{V}_{t+1}(x_{it+1}, \omega_{mt+1})|a_{it}, x_{it}, \omega_{mt})\big]\Big\} \end{aligned} \tag{12}$$

Using expectational errors to eliminate the expecation, this can now be rewritten as:

$$\begin{aligned} \ln\left(\frac{p_m(j_{it}|k_{it}, \omega_{mt})}{p_m(a_{it}|k_{it}, \omega_{mt})}\right) = & \ \big(\pi(j_{it}, x_{it}, w_{mt}) - \pi(a_{it}, x_{it}, w_{mt})\big) \\ & + \beta\Big\{\bar{V}_{mt+1}(k_{it+1}(j_{it}), \omega_{mt+1}) - \bar{V}_{mt+1}(k_{it+1}(a_{it}), \omega_{mt+1})\Big\} \\ & + \beta\Big\{e^V(k_{it+1}(j_{it}), \omega_{mt}, \omega_{mt+1}) - e^V(k_{it+1}(a_{it}), \omega_{mt}, \omega_{mt+1})\Big\} \end{aligned} \tag{13}$$

Using equation 9, we can write the difference in ex ante value functions in equation 13 in terms of conditional value functions and conditional choice probabilities with respect to some arbitrary choice l:

$$\begin{aligned} & \bar{V}_{mt+1}(k_{it+1}(j_{it}), \omega_{mt+1}) - \bar{V}_{mt+1}(k_{it+1}(a_{it}), \omega_{mt+1}) = \\ & \qquad v_m(l_{it+1}, j_{it}, \omega_{mt+1}) - v_m(l_{it+1}, a_{it}, \omega_{mt+1}) \\ & \qquad - \ln\big(p_m(l_{it+1}|j_{it}, \omega_{mt+1})\big) + \ln\big(p_m(l_{it+1}|a_{it}, \omega_{mt+1})\big) \end{aligned} \tag{14}$$

Due to one-period finite dependence, continuation values within $v_m(l_{it+1}, j_{it}, \omega_{mt+1})$ and $v_m(l_{it+1}, a_{it}, \omega_{mt+1})$ are identical; therefore the first-term on the right-hand side of equation 14 can be written as a difference in next period profits:

$$\begin{aligned} & \bar{V}_{mt+1}(k_{it+1}(j_{it}), \omega_{mt+1}) - \bar{V}_{mt+1}(k_{it+1}(a_{it}), \omega_{mt+1}) = \\ & \qquad \pi_m(l_{it+1}, j_{it}, \omega_{mt+1}) - \pi_m(l_{it+1}, a_{it}, \omega_{mt+1}) \\ & \qquad - \ln\big(p_m(l_{it+1}|j_{it}, \omega_{mt+1})\big) + \ln\big(p_m(l_{it+1}|a_{it}, \omega_{mt+1})\big) \end{aligned} \tag{15}$$

Now, substituting equation 15 into equation 13, and setting $l = j$:

$$\ln\left(\frac{p_m(j_{it}|k_{it},\omega_{mt})}{p_m(a_{it}|k_{it},\omega_{mt})}\right)+\beta\ln\left(\frac{p_m(j_{it}|j_{it},\omega_{mt+1})}{p_m(j_{it}|a_{it},\omega_{mt+1})}\right) \tag{16}$$
$$= \big(\pi_m(j_{it},k_{it},\omega_{mt}) - \pi_m(a_{it},k_{it},\omega_{mt})\big)$$
$$+ \beta\big(\pi_m(j_{it},j_{it},\omega_{mt}) - \pi_m(j_{it},a_{it},\omega_{mt})\big)$$
$$+ \beta\Big\{e^V(k_{it+1}(j_{it}),\omega_{mt},\omega_{mt+1}) - e^V(k_{it+1}(a_{it}),\omega_{mt},\omega_{mt+1})\Big\}$$

Finally, we can use the definition of $\pi_m(\cdot)$, as well as the specifications of rents and conversion costs to arrive at the estimating equation:

$$Y_{jmt}(k) = (\eta^0_{jk} - \eta^0_{ak}) + \sum_{\bar{q}=2}^{6} \mathbb{1}(q_m = \bar{q})\Big(\eta^{\bar{q}}_{jk} - \eta^{\bar{q}}_{ak}\Big) \tag{17}$$
$$+ (\theta^0_{jk}P_{jct} - \theta^0_{ak}P_{act}) + \sum_{\bar{q}=2}^{6} \mathbb{1}(q_m = \bar{q})\Big(\theta^{\bar{q}}_{jk}P_{jct} - \theta^{\bar{q}}_{ak}P_{act}\Big)$$
$$+ \beta\Big\{(\eta^0_{jj} - \eta^0_{ja}) + \sum_{\bar{q}=2}^{6} \mathbb{1}(q_m = \bar{q})\Big(\eta^{\bar{q}}_{jj} - \eta^{\bar{q}}_{ja}\Big)$$
$$+ (\theta^0_{jj} - \theta^0_{ja})P_{jct} + \sum_{\bar{q}=2}^{6} \mathbb{1}(q_m = \bar{q})\Big(\theta^{\bar{q}}_{jj} - \theta^{\bar{q}}_{ja}\Big)P_{jct}\Big)\Big\} + u_{jmt}$$

where:

$$Y_{jmt}(k) \equiv \ln\left(\frac{p_m(j_{it}|k_{it},\omega_{mt})}{p_m(a_{it}|k_{it},\omega_{mt})}\right) + \beta\ln\left(\frac{p_m(j_{it}|j_{it},\omega_{mt+1})}{p_m(j_{it}|a_{it},\omega_{mt+1})}\right) \tag{18}$$

$$u_{jmt} \equiv \Big(\xi(j_{it},k_{it},\omega_{mt}) - \xi(a_{it},k_{it},\omega_{mt})\Big) \tag{19}$$
$$+ \beta\Big\{e^V(k_{it+1}(j_{it}),\omega_{mt},\omega_{mt+1}) - e^V(k_{it+1}(a_{it}),\omega_{mt},\omega_{mt+1})\Big\} + u_{jmt}$$

1.4 Identification

1.5 Estimation

The dependent variable Y_{jit} in equation **??** is a function of conditional choice probabilities $p_j(x_{it},\omega_{mt})$, which—while not directly observed—can be estimated in a first-stage. Specifically, in order to construct Y_{jit}, we estimate:

$$p_j\left(x_{it}(k_{it}),\omega_{mt}\right) = F\left(\gamma^j_0 + \gamma^j_1(P^j_{ct} - P^k_{ct}) + \gamma^j_2 LCC_i + \theta^j_3(P^j_{ct} - P^k_{ct}) \times LCC_i + \delta_c + \mu_t\right) \tag{20}$$

where $F(\cdot)$ is the cumulative distribution function of a logistic distribution. Estimates of equation 23 measure landowner responses to change in land use returns without taking into account dynamic behavior. Thus they are both necessary inputs in constructing the dependent variable for our dynamic discrete choice estimator, and they provide a basis of comparison for estimates from the dynamic discrete choice estimator. Based on our estimate of equation 23, we predict conditional choice probabilities for current period choices j and a, and next-period choices of J conditional on choosing j and a in the present period. We uses these conditional choice probabilities to calculate

References

Arcidiacono, Peter and Paul B Ellickson, "Practical methods for estimation of dynamic discrete choice models," *Annu. Rev. Econ.*, 2011, *3* (1), 363–394.

_ **and Robert A Miller**, "Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity," *Econometrica*, 2011, *79* (6), 1823–1867.

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

Kalouptsidi, Myrto, Paul T Scott, and Eduardo Souza-Rodrigues, "Linear IV regression estimators for structural dynamic discrete choice models," *Journal of Econometrics*, 2020.

Lubowski, Ruben N, Andrew J Plantinga, and Robert N Stavins, "Land-use change and carbon sinks: econometric estimation of the carbon sequestration supply function," *Journal of Environmental Economics and Management*, 2006, *51* (2), 135–152.