Supplement to "Identification of Counterfactuals in Dynamic Discrete Choice Models

Myrto Kalouptsidi[*], Paul T. Scott[†], Eduardo Souza-Rodrigues[‡]

Harvard University, CEPR and NBER, NYU Stern School of Business, University of Toronto

February 2020

This supplemental material consists of the following sections: Section B presents the data sources, explains the construction of the variables used in the empirical application, and shows some summary statistics. Section C discusses the implementation of the empirical exercise based on a dynamic model of farmers land use decisions.

B Data and Summary Statistics

Table B1 lists our data sources. All are publicly available for download save DataQuick's land values. Our main sample is based on a sub-grid of the Cropland Data Layer (CDL), a high-resolution (30-56m) annual land-use data that covers the entire contiguous United States since 2008. We took a 840m sub-grid of the CDL within those counties appearing in our DataQuick database.[1] DataQuick collects transaction data from about 85% of US counties and reports the associated price, acreage, parties involved, field address and other characteristics. The coordinates of the centroids of transacted parcels in the DataQuick database are known. To assign transacted parcels a land use, we associate a parcel with the nearest point in the CDL grid.

A total of 91,198 farms were transacted between 2008 to 2013 based on DataQuick. However, we dropped non-standard transactions and outliers from the data. First, because we are interested in the agricultural value of land (not residential value), we only consider transactions of parcels for

[*]Department of Economics, Harvard University, Littauer Center, Cambridge, MA 02138, myrto@fas.harvard.edu

[†]Stern School of Business, New York University, Kaufman Management Center, 44 W. 4th St., New York, NY 10012, ptscott@stern.nyu.edu.

[‡]Department of Economics, University of Toronto, Max Gluskin House, 150 St. George St., Toronto, Ontario M5S 3G7, Canada, e.souzarodrigues@utoronto.ca

[1]The 840m grid scale was chosen for two reasons. First, it provides comprehensive coverage (i.e., most large agricultural fields are sampled) without providing too many repeated points within any given parcel. Second, the CDL data changed from a 56m to a 30m grid, and the 840 grid size allows us to match points across years where the grid size changed while matching centers of pixels within 1m of each other. The CDL features crop-level land cover information. See Scott (2013) for how "crops" and "non-crops" are defined.

Table B1: Data Sources

Dataset	Description	Source
Cropland Data Layer	Land cover	http://nassgeodata.gmu.edu/CropScape/
DataQuick	Real estate transactions, assessments	DataQuick
US Counties	County boundaries	http://www.census.gov/cgi-bin/geo/shapefiles2010/layers.cgi
GAEZ Database	Protected land, soil type	http://www.gaez.iiasa.ac.at/
SRTM	Topographical – altitude and slope	http://dds.cr.usgs.gov/srtm/
NASS Quick Stats	Yields, prices, pasture rental rates	http://www.nass.usda.gov/Quick_{}{}{}{}{}{}Stats/
ERS	Operating costs	http://www.ers.usda.gov/data-products/commodity-costs-and-returns.aspx
NOOA Urban Centers	Urban center locations and populations	%{}{}{}{}{} http://www.nws.noaa.gov/geodata/catalog/national/html/urban.

which the municipal assessment assigned zero value to buildings and structures. Additionally, we drop transactions featuring multi-parcels, transactions between family members, properties held in trust, and properties owned by companies. Finally, we drop transactions with extreme prices: those with value per acre greater than $50,000, total transaction price greater than $10,000,000, or total transaction price less than $60; these are considered measurement error. After applying the selection criteria, there remained $24,643$ observations (transactions).

To obtain a rich set of field characteristics, we use soil categories from the Global Agro-Ecological Zones database and information on protected land from the World Database on Protected Areas. Protected land was dropped from all analyses. The NASA's Shuttle Radar Topography Mission (SRTM) database provides detailed topographical information. The raw data consist of high-resolution (approx. 30m) altitudes, from which we calculated slope and aspect, all important determinants of how land is used. Characteristics such as slopes and soil categories are assigned to fields/parcels using nearest neighbor interpolation.

To derive a measure of nearby developed property values, we find the five restaurants nearest to a field, and we average their appraised property values. For each field, we also compute the distance to the nearest urban center with a population of at least 100,000. Location of urban centers were obtained from the National Oceanic and Atmospheric Administration (NOAA).

Finally, we use various public databases on agricultural production and costs from the USDA. The final dataset goes from 2010 to 2013 for 515 counties and from 2008 to 2013 for 132 counties. Crop returns are based on information on yields, prices received, and operating expenditures; non-crop returns are based on more sparse information on pasture land rental rates (see Scott (2013)).

Table B2 presents some summary statistics. Table B3 compares the transacted fields (in DataQuick) to all US fields (in the CDL). Overall, the two sets of fields look similar. In particular, the probability of keeping (switching to) crops is very similar across the two datasets.

Table B2: Summary Statistics

Statistics	Mean	Std Dev	Min	Max
In Cropland	0.147	0.354	0	1
Switch to Crops	0.0162	0.126	0	1
Keep Crops	0.849	0.358	0	1
Crop Returns ($)	228	112	43	701
Slope (grade)	0.049	0.063	0	0.702
Altitude (m)	371	497	−6	3514
Distance to Urban Center (km)	79.8	63.7	1.22	362
Nearest commercial land value ($/acre)	159000	792000	738	73369656
Land value ($/acre)	7940	9720	5.23	50000

A slope of 1 refers to a perfect incline and a slope of 0 refers to perfectly horizontal land.

Table B3: Dataquick vs CDL Data – Time Invariant Characteristics

Mean by dataset	DataQuick	CDL
In Cropland	0.147	0.136
Switch to Crops	0.0162	0.0123
Keep Crops	0.849	0.824
Crop Returns ($)	228	241
Slope (grade)	0.049	0.078
Altitude (m)	371	688
Distance to Urban Center (km)	79.8	103
Nearest commercial land value ($/acre)	159000	168000

C Dynamic Land Use Model and Estimation

C.1 Model with Unobserved States

As mentioned in the main text, we augment the empirical model by allowing for unobserved market states, following Scott (2013). The per period payoff becomes:

$$\pi\left(a, k_{imt}, w_{mt}, s_{im}, \varepsilon_{imt}\right) = \theta_0\left(a, k_{imt}, s_{im}\right) + \theta_1 Z\left(a, w_{mt}\right) + \xi\left(a, k_{imt}, w_{mt}, s_{im}\right) + \varepsilon_{imt} \tag{C1}$$

where $\xi\left(a, k, w, s\right)$ captures unobservable variation in returns, and the idiosyncratic shock ε_{it} has a logistic distribution. (Without loss of generality, $\xi\left(a, k, w, s\right)$ is mean-zero for all $\left(a, k, w, s\right)$.) We construct returns $Z^a_{mt} \equiv Z\left(a, w_{mt}\right)$ using county-year information (expected prices and realized yields for major US crops, as well as USDA cost estimates) as in Scott (2013).[2] As described below, identification requires exclusion restrictions on $\xi\left(a, k, w, s\right)$ (see also Kalouptsidi, Scott, and Souza-Rodrigues, 2018).

C.2 Payoff Parameter Estimation

Throughout this section, we use t-subscripts in place of explicitly writing the aggregate state variable w_{mt}. We also omit the subscripts i (fields) and m (counties) to simplify notation. The derivation relies on two crucial assumptions: (a) agents are small; i.e., changing the action of any agent at time t does not affect the distribution of w_{t+1}, and (b) agents have rational expectations.

Here, we consider two estimators for the payoff function. Let $p^c_t\left(k, s\right)$ denote the probability of choosing action "crops" at time period t given state k for a field of type s, and let σ be the scale parameter of the logit shocks (we discuss this further below). We begin with Scott's (2013) linear estimating equation for a dynamic model with logit errors; we refer the interested reader to Scott (2013) (see also Kalouptsidi, Scott, and Souza-Rodrigues, 2018) for the derivation of the following equation:

$$Y_t\left(k, s\right) = \tilde{\theta}_0\left(k, s\right) + \theta_1\left(Z_t\left(c, s\right) - Z_t\left(nc, s\right)\right) + \tilde{\xi}_{k,s,t} + \tilde{e}_{k,s,t} \tag{C2}$$

[2] We refer the interested reader to Scott (2013) for details of constructing the measure of observed returns Z. Due to data limitations, we restrict Z to depend only on $\left(a, w_{mt}\right)$. One important difference from Scott (2013) is that we have field level observable characteristics s_{im} and they affect land use switching costs.

where

$$\begin{aligned}
Y_t(k,s) &\equiv \ln\left(\frac{p_t^c(k,s)}{1-p_t^c(k,s)}\right)+\beta\ln\left(\frac{p_{t+1}^c(0,s)}{p_{t+1}^c(k'(nc,k),s)}\right)\\
\tilde{\theta}_0(k,s) &\equiv \left(\theta_0(c,k,s)-\theta_0(nc,k,s)\right)/\sigma\\
&\quad +\beta\left(\theta_0(c,0,s)-\theta_0(c,k'(nc,k),s)\right)/\sigma\\
\theta_1 &\equiv 1/\sigma\\
\tilde{\xi}_{k,s,t} &\equiv \xi_t(c,k,s)-\xi_t(nc,k,s)\\
&\quad +\beta\left(\xi_{t+1}(c,0,s)-\xi_{t+1}(c,k'(nc,k),s)\right)
\end{aligned}$$

$$\begin{aligned}
\tilde{e}_{k,s,t} &\equiv \beta\left(E_t\left[V_t(0,s)\right]-V_t(0,s)\right)\\
&\quad -\beta\left(E_t\left[V_{t+1}(k'(nc,k),s)\right]-V_{t+1}(k'(nc,k),s)\right).
\end{aligned}$$

Ultimately, this is a linear equation that can be used to estimate the parameters of the payoff function with no need to solve the agent's dynamic optimization problem.

On the left hand side of equation (C2), we have a dependent variable which is a function of conditional choice probabilities (which are estimated in a first stage, described below in Section C.3) and the discount factor (which is imputed; we assume it equals 0.95).

On the right hand side of (C2), the intercept term $\tilde{\theta}_0$ is a combination of intercepts of the payoff function θ_0. We discuss the identification of θ_0 in more detail below, for this is essentially where the two estimators differ.

The error term has two components, $\tilde{\xi}$ and $\tilde{e}$. The term $\tilde{\xi}$ is a function of ξ, representing unobservable variation in returns, while $\tilde{e}$ is a function of expectational error terms. Because Z and ξ may be correlated, we follow Scott (2013) and implement an instrumental variable estimator. To do so, we need exclusion restrictions of the form

$$E\left[\nu_{k,s,t}\left(\tilde{\xi}_{k,s,t}+\tilde{e}_{k,s,t}\right)\right]=0, \tag{C3}$$

where $\nu_{k,s,t}$ is a vector of instrumental variables. Given that agents have rational expectations, $\tilde{e}_{k,s,t}$ is uncorrelated with any function of variables in the time-t information set by construction. For this reason, $E\left[\nu_{k,s,t}\tilde{e}_{k,s,t}\right]=0$ holds for any $\nu_{k,s,t}$ in the time-t information set and the question of whether equation (C3) is valid becomes a question of whether $E[\nu_{k,s,t}\tilde{\xi}_{k,s,t}]=0$. Such a restriction is a substantive assumption as exclusion restrictions for instrumental variables typically are.[3]

We take first-differences for each field and field state, implicitly allowing for $\tilde{\xi}_{k,s,t}$ to have fixed effects for s and k (interacted).[4] After taking first differences, the instruments we use are: a

[3]If we were willing to assume that $E[(Z_{s,t}^c-Z_{s,t}^{nc})\tilde{\xi}_{k,s,t}]=0$, then we could estimate equation (C2) using ordinary least squares.

[4]Note that we predict CCPs for each field state k, not just for the field state actually observed on the field, so we can take these first differences for each k regardless of the actual path of k for the field.

constant term, caloric yields, and the lagged value of $Z^{c}_{s,t} - Z^{nc}_{s,t}$.[5] The moment restrictions are used to estimate θ_1. We then form estimates of $\tilde{\theta}_0(k,s)$ by averaging over the residuals for each (k,s) pair.

Up to this point, our two estimators coincide; i.e., our two estimators agree on the estimates of θ_1 and $\tilde{\theta}_0(k,s)$. The estimators differ when it comes from mapping the estimates of $\tilde{\theta}_0(k,s)$ to estimates of $\theta_0(\cdot,k,s)$. Notice that for each type s, equation (C2) includes one intercept parameter $\tilde{\theta}_0(k,s)$ for each field state k. However, the original payoff function involves two intercept parameters ($\theta_0(c,k,s)$ and $\theta_0(nc,k,s)$) for each (s,k) combination. Hence, the need for restrictions for the identification of the model (and our claim in Section 3.5 that θ_0 is not identified without restrictions).

Our first estimator (the CCP estimator) imposes the following restrictions on θ_0:

$$\forall k,s: \quad \theta_0(nc,k,s) = 0. \tag{C4}$$

After imposing (C4), we can solve for $\theta_0(c,k,s)$ from our $\tilde{\theta}_0(k,s)$ estimates, recalling that

$$\tilde{\theta}_0(k,s) \equiv (\theta_0(c,k,s) - \theta_0(nc,k,s))/\sigma + \beta(\theta_0(c,0,s) - \theta_0(c,k'(nc,k),s))/\sigma, \tag{C5}$$

noting that equations (C4) and (C5) represent six linearly independent equations in six unknowns for each (k,s) pair. (And noting that the scale parameter σ is identified given that $\theta_1 \equiv 1/\sigma$.)

Our second estimator (the V-CCP estimator) does not impose equation (C4), and instead uses additional information in resale prices. In order to relate observed resale prices to farmer's payoff and value functions, we need a model of transaction prices. We assume that resale prices measure farmer's ex-ante value functions; i.e.,

$$\ln p^{RS}_t = \ln\left(\tilde{V}_t(k,s)\right) + \eta_t, \tag{C6}$$

where p^{RS}_t is the resale price of a field, η_t is measurement error, and we will explain the reason for the tilde on the value function below. Using resale prices as signals of the value function can be justified by assuming that there is a competitive market for buying farms – see Kalouptsidi (2014) for further discussion of this assumption in the context of bulk shipping.

We estimate a flexible model of how resale prices depend on (k,s,t), much like Kalouptsidi (2014) (see Section C.4 for details about the implementation). Fitted values from this regression can be used as estimates of the value function, but an important caveat is that we must consider the scale of the utility function when interpreting the estimates. In econometric discrete choice models, we typically impose a scale normalization on the model that sets the variance of the idiosyncratic shocks equal to a convenient number (e.g., unity for a probit model of $\pi^2/6$ for a logit model). In our parametric land use model, the coefficient on returns, θ_1, reflects this normalization: the parameter

[5]See Scott (2013) for the measurement of caloric yields.

θ_1 can be understood as the scalar we need to multiply by to convert the units from dollars to utils. When we estimate a hedonic model of the value function, the value function is measured in dollars. Therefore, to convert from the estimated value function to the scale-normalized value function we should multiply by θ_1:

$$V_t(k,s) = \theta_1 \tilde{V}_t(k_{it}, s_{it}).$$

A relationship between value functions and the payoff function can be derived as follows:

$$\begin{aligned} V_t(k,s) &= v_t(c,k,s) + \psi_c(p_t^c(k,s)) \\ &= \pi_t(c,k,s) + \beta E_t[V_{t+1}(k'(c,k),s)] + \psi_c(p_t^c(k,s)) \\ &= \pi_t(c,k,s) + \beta V_{t+1}(k'(c,k),s) + e_{k,s,t} + \psi_c(p_t^c(k,s)) \end{aligned}$$

where

$$e_{k,s,t} \equiv \beta\left(E_t[V_t(k'(c,k),s)] - V(k'(c,k),s)\right).$$

Ultimately, we can write the payoff function as a function of conditional choice probabilities (estimated in a first stage), value functions (estimated using retail prices in a first stage), and an expectational error term (mean zero):

$$\pi_t(c,k,s) = V_t(k,s) - \beta V_{t+1}(k'(c,k),s) - \psi_c(p_t^c(k,s)) - e_{k,s,t}. \tag{C7}$$

Recalling that the measured version of the value function needs to be converted from dollars to utils to be on the same scale as the normalized payoff function, we have

$$\pi_t(c,k,s) = \theta_1\left(\tilde{V}_t(k,s) - \beta\tilde{V}_{t+1}(k'(c,k),s)\right) - \psi_c(p_t^c(k,s)) - e_{k,s,t}. \tag{C8}$$

Noting that an estimate of θ_1 can be obtained from the CCP estimator, we can then obtain estimates of payoffs using equation (C8), simply by plugging in the estimated values of θ_1, $\tilde{V}$ and p.[6] More to the point, we can obtain estimates of the intercept parameters:

$$\theta_0(c,k,s) = -\theta_1 Z_t(c,s) + \theta_1\left(\tilde{V}_t(k,s) - \beta\tilde{V}_{t+1}(k'(c,k),s)\right) - \psi_c(p_t^c(k,s)) - e_{k,s,t}. \tag{C9}$$

The V-CCP estimator uses equation (C9) to estimate $\theta_0(c,k,s)$ by averaging the right-hand-side of (C9) over time. Finally, the estimates of $\theta_0(nc,k,s)$ are then recovered from equation (C5).

Note that we could alternatively estimate $\theta_0(nc,k,s)$ from an equation like (C9), but using non-

[6] Recall that estimating θ_1 with the CCP estimator does not require any identifying restrictions on θ_0. Consider equation (C2), a regression equation that allows us to estimate θ_1 and $\tilde{\theta}_0$. The identifying restrictions are only needed if we want to map from $\tilde{\theta}_0$ to θ_0.

crops as the action instead of crops. Thus, we have over-identifying restrictions. As the primary reason we consider the V-CCP estimator is to replace the *a priori* identifying restrictions in the CCP estimator with a more data-driven approach, we only take as much information as we need from the resale prices to fully identify the payoff function. If we were to use more information from the resale prices, then the two estimators might not agree on the value of $\tilde{\theta}_0(k,s)$, an object that is identified from CCP data without imposing identifying restrictions. Our two estimators only differ when it comes to parameters that cannot be identified from CCP data without restrictions. Thus, by comparing these two estimators, we isolate the impact of identifying restrictions.

C.3 Conditional Choice Probabilities

We estimate conditional choice probabilities using a semiparametric logit model. The model is fully flexible over field states and year, but smooth across counties. In particular, we maximize the following log likelihood function:

$$\max_{\theta_{ckt}} \sum_{m' \in S_m} \sum_{i \in I_{m'}} w_{m,m'} I\left[k_{imt} = k\right] \left\{ \begin{array}{c} I\left[a_{imt} = c\right] \log\left(p_{mt}\left(c, k, s_{im}; \theta_{ckt}\right)\right) \\ +I\left[a_{imt} = nc\right] \log\left(1 - p_{mt}\left(c, k, s_{im}; \theta_{ckt}\right)\right) \end{array} \right\}$$

where S_m is the set of counties in the same US state as m, I_m is the set of fields in county m, $w_{m,m'}$ is the inverse squared distance between counties m and m', and $I[.]$ is the indicator function. The conditional choice probability is parameterized as follows:

$$p_{mt}\left(c, k, s_{im}; \theta_{ckt}\right) = \frac{\exp\left(s'_{im}\theta_{ckt}\right)}{1 + \exp\left(s'_{im}\theta_{ckt}\right)}.$$

Note that without fields' observable characteristics, this regression would amount to taking frequency estimates for each county, field state, and year, with some smoothing across counties. Including covariates allows for within-county field heterogeneity. The final specification for the conditional choice probabilities only uses $slope_{im}$ among regressors because it proved to be the most powerful predictor of agricultural land use decisions (after conditioning on county and field state).

The set of counties in S_m only includes counties which also appear in the DataQuick database. For some states, the database includes a small number of counties, so in these cases we group two or three states together. For example, only one county in North Dakota appears in our sample, and it is a county on the eastern border of North Dakota, so we combine North Dakota and Minnesota. Thus, for each county m in North Dakota or Minnesota, S_m represents all counties in both states in our sample.[7]

[7] In particular, we form a number of groups for such cases: Delaware and Maryland; North Dakota and Minnesota; Idaho and Montana; Arkansas, Louisiana, and Mississippi; Kentucky and Ohio; Illinois, Indiana, and Wisconsin; Nebraska and Iowa; Oregon and Washington; Colorado and Wyoming.

Table C1: Land Resale Price Regression

VARIABLES	(1) log(land value)
log(distance to urban center)	-0.471***
	(0.0297)
commercial land value	0.102***
	(0.00930)
slope	-1.654***
	(0.160)
alt	-0.000226**
	(9.00e-05)
Observations	24,643
R-squared	0.318

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Ommitted: soil, county, year, and field state dummies

as well as interactions with returns.

For the sake of precision, rather than only estimating CCPs using the CDL sample that was merged with resale data, we used the full 840m sub-grid of fields from the CDL (848,384 fields) for the CCP estimation. We then predicted CCPs and estimated payoff functions using fields that were merged with the resale data.

C.4 Resale Price Regression

Next, we discuss how we estimate the value function from resale prices. We view that our resale market assumptions are not overly restrictive in the context of rural land which features a large number of small agents. The land resale market is arguably thick, with a large number of transactions taking place every year.[8] Moreover, we are able to control for a rich set of field characteristics. Finally, we did not find evidence of selection on land use changes upon resale, as discussed below.

As our transaction data is much more sparse than our choice data, we adopt a more restrictive (parametric) form for modeling land values. We estimate the following regression equation:

$$\ln p_{it}^{RS} = X_{it}'\theta_V + \eta_{it}, \tag{C10}$$

where p_{it}^{RS} is a transaction price (in dollars per acre), and X_{it} is a vector of characteristics for the corresponding field. The covariates X_{it} include all variables in Table B2 (i.e. k, slope, altitude, distance to urban centers, nearby commercial values). They also include year dummies, returns interacted with year dummies, field state dummies interacted with year dummies, and county

[8] Comparing DataQuick with the CDL data we see that 1.4–2% of fields are resold every year. Moreover, the USDA reports that in Wisconsin there are approximately 100 thousand acres transacted every year (about 1000 transactions) out of 14.5 million acres of farmland (seemingly information on other states is not available).

Table C2: Land use and transactions

VARIABLES	(1) incrops2010	(2) incrops2011	(3) incrops2012	(4) incrops2013
soldin2009	0.000647			
	(0.00604)			
soldin2010	0.000116	0.00364		
	(0.00326)	(0.00334)		
soldin2011	-0.00117	0.000629	-0.00159	
	(0.00316)	(0.00324)	(0.00330)	
soldin2012		-0.000620	-0.00472	0.00411
		(0.00306)	(0.00313)	(0.00265)
soldin2013			-0.00962***	-0.000445
			(0.00306)	(0.00256)
Observations	23,492	23,492	23,492	23,492
R-squared	0.666	0.698	0.717	0.757

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Linear probability model. Omitted covariates include current returns, field state, US state, slope, local commercial land value, distance to nearest urban center, and interactions.

dummies.

Table C1 presents the estimated coefficients. Although not shown in the table, the estimated coefficients of k are significant and have the expected signs (the large number of interactions makes it difficult to add them all in the table). This is important for the second stage estimation, as k is the main state variable included in the switching cost parameters $\theta_0(a, k)$.

Note that, because field acreage is available only in the DataQuick dataset, when merging with the CDL and remaining datasets we lose this information. This implies, for example, that acreage cannot be a covariate in the choice probabilities. For this reason, we choose a specification for the value function that regresses price per acre on covariates. The value of our R^2 in our regression is a direct consequence of this fact. When we use total land prices as the dependent variable and include acres on the covariates we obtain R^2 as high as 0.8.

Finally, we briefly discuss the possibility of selection on transacted fields. As shown previously in Table B3 of Section B, the characteristics of the transacted fields (in DataQuick) look similar to all US fields (in the CDL). Furthermore, we investigate whether land use changes upon resale. Using a linear probability model we find no such evidence (see Table C2). We regress the land use decision on dummy variables for whether the field was sold in the current, previous, or following year as well as various control variables. In regressions within each cross section, ten of the eleven coefficients on the land transaction dummy variables are statistically insignificant, and the estimated effect on the probability of crops is always less than 1% (see Table C2). We have tried alternative specifications such as modifying the definition of the year to span the planting year rather than calendar year, and yet we have found no evidence indicating that there is an important connection between land transactions and land use decisions.

References

KALOUPTSIDI, M. (2014): "Time to build and fluctuations in bulk shipping," *The American Economic Review*, 104(2), 564–608. C.2

KALOUPTSIDI, M., P. T. SCOTT, AND E. A. SOUZA-RODRIGUES (2018): "Linear IV Regression Estimators for Structural Dynamic Discrete Choice Models," *NBER Working Paper 25134*. C.1, C.2

SCOTT, P. T. (2013): "Dynamic Discrete Choice Estimation of Agricultural Land Use," *Working Paper*. 1, B, C.1, C.1, C.2, 2, C.2, 5