A dynamic discrete choice model of land use conversion: Data

July 19, 2021

1 Land use returns

1.1 Crop returns

We construct annual county-specific crop rents based on weighted averages of returns in each county for eleven principal crops that account for 93% of harvested cropland (excluding hay) in the US according to the 2017 Census of Agriculture.[1] County-level crop-specific returns were estimated according to:

$$R_{nct} = (P_{n,s(c),t} - C_{n,r(c),t})Yield_{nct} \tag{1}$$

where R_{nct} represents returns per acre for crop n in county c in year t, $P_{n,s(c),t}$ represents the per unit prices for crop n in state $s(c)$ in year t, and $C_{n,r(c),t}$ represents the per unit production costs for crop n in Farm Resource Region $r(c)$ in year t, and $Yield_{nct}$ represents crop yield per acre for crop n in county c in year t. State-level crop prices were collected from... Data on costs per planted acre were based on operating costs observed at the Farm Resource Region-level, collected from Economic Research Service (ERS) Commodity Costs and Returns.

We multiplied net returns per unit of production $(P_{n,s(c),t} - C_{n,r(c),t})$ by average production per acre (yield) in each county to provide a county-specific measure of net returns per acre for production of each crop. Yield data were collected from the Census of Agriculture for all census years in the time span of our data (2002, 2007, 2012), and yields in intervening years were obtained by linearly interpolating between census years at the county-level (TRUE?). For privacy purposes, the National Agricultural Statistics Service withholds data when there are few producers in a a given reporting area; therefore, we do not observe yields for every county and crop in every year. For counties and crops with missing yields, we interpolate based on yields for the same crop and year in other counties within the same Farm Resource Region.[2]

---

[1] The eleven crops are corn, sorghum, soybeans, winter wheat, durum wheat, other spring wheat, barley, oats, rice, upland cotton, and pima cotton. These are the same eleven crops used by **?** to construct his measure of crop returns.

[2] Our interpolation method is similar to the one used to smooth conditional choice probabilities, based on the method used by **?**. For each pair of counties c and c' in the same Farm Resource Region, we calculate weights $w_{cc'} = (1 + d_{cc'})^{-2}$ where $d_{cc'}$ is the distance between counties c and c'. For counties with missing yield data, we

After estimating crop-specific net returns per acre for each county, we calculate overall cropland returns as a weighted average of crop-specific returns, where weights are based on acreages in each crop in state $s(c)$. Specifically, we calculate:

$$R_{ct} = \frac{\sum_{n \in \mathcal{N}} A_{n,s(c),t} R_{nct}}{\sum_{n \in \mathcal{N}} A_{n,s(c),t}} \tag{3}$$

where N is the set of eleven primary crops and $A_{n,s(c),t}$ is total acres planted in crop n in state $s(c)$ in year t according to the Census of Agriculture (OR THE SURVEY)? We weight crop-specific returns by state-level acreages since many crop-specific acreages are withheld for privacy purposes in many individual counties. In some states, no single county reports acres planted for a given crop. For counties in these states, we calculate R_{ct} using weights based on total acres planted in a given crop within the same Farm Resource Region.

We calculate R_{ct}^G, average returns to cropland in county c in year t, inclusive of government payments, by adding government payments per acre to R_{ct}. Formally, we define:

$$R_{ct}^G \equiv R_{ct} + \frac{G_{ct}}{\tilde{A}_{c,\tilde{t}(t)}} \tag{4}$$

where G_{ct} is NEED DETAILS ON WHAT DATA WE'RE USING FOR G and $\tilde{A}_{c,\tilde{t}(t)}$ is the number of acres in cropland in county c, as measured by the NRI in the nearest NRI year $\tilde{t}(t)$ to year t. Returns to cropland are summarized in Figure **??**. Panel A maps geographic variation in crop returns across the US in 2002, the beginning of the sample period. Panel B is a histogram illustrating the distribution crop returns across counties in 2002. Panel C shows how average crop returns (weighted by total county cropland area) vary from 2002-2012.

1.2 Pasture returns

Estimates of pasture rents come from the National Agricultural Statistics Service (NASS), who have provided data on per-acre rents on pasturelands since 1994. The most precise geographic level at which pasture rents are available varies by year, however. For each NRI year, we collected pasture rents at the most precise geographic level available from NASS. As well, we used 2007 pasture rent data in place of 2008 data since the 2007 data were available at a higher geographic resolution. For each year, the primary geographic level over which we collected pasture rents was: 1997 (state), 2002 (state), 2007 (Agricultural Statistical District), 2012 (county), 2015 (state). Agricultural Statistical Districts are generally sub-state geographies defined by NASS as groupings of counties in each state according to geography, climate, and cropping practices, with typically 5–10 districts

---

estimate yield as:

$$\widehat{Yield}_{nct} = \frac{\sum_{c' \in \mathcal{C}_r(c)} w_{cc'} Yield_{nct}}{\sum_{c' \in \mathcal{C}_r(c)} w_{cc'} \mathbb{1}(Yield_{nct} > 0)} \tag{2}$$

where $\mathcal{C}_{r(c)}$ is the set of counties in the same Farm Resource Region r as county c and $Yield_{nct}$ is set to be zero in counties in which yield is missing.

per state. Pasture rents were sometimes available only for more aggregated geographies than are listed here, and in rare cases pasture rents were available at less aggregated geographies. In all cases, we used the pasture rents for the least aggregated geography available.

1.3 Conservation Reserve Program returns

We estimate rents on CRP land based on Conservation Reserve Program statistics reported by the US Department of Agriculture Farm Service Agency (FSA). The FSA provides historical county-level data on CRP enrollment and average rental payments per acre from 1986 to present. We adjust reported average per acre rental payments in NRI years 1987–2015 to 2010 dollars and use these as a measure of rents on CRP land.

1.4 Forest returns

County-level estimates of net returns from forestland for NRI years 2002–2012 were provided by Mihiar and Lewis (2019). They construct their measure of forestland returns based on the estimated net present value of a an acre of bare forest land in each county. They collect county-by-species stumpage prices from miscellaneous sources documented in Mihiar and Lewis (2017) and average harvest ages based on USDA Forest Inventory Analysis data. Using these data together with estimated county-by-species timber growth equations, they estimate calculate species-specific one-rotation profits for each county. They then annualize these estimates, convert them to 2010 USD, and construct a weighted average of forest net returns in each county, where weights are based on the share of timber production due to each species. We use this weighted average as our measure of forest net returns. More details on the construction of forest net returns are available in Mihiar and Lewis (2019).

1.5 Rangeland returns

1.6 Urban returns

We estimate urban returns by adapting the strategy used by Lubowski et al. (2006) and more recently by Mihiar and Lewis (2019). From US Census Public Use Microdata Sample (PUMS) data, we collect Public Use Microdata Area (PUMA)-level property value data. Because we are interested in potential rents from converting into urban use, and rents on newly converted land may differ from rents on land that was converted into urban use long ago, we calculate median property values within each PUMA among only properties built in the previous five years. We convert PUMA-level median property values to county-level median property values by assuming that properties are uniformly distributed within Census tracts, and calculating county-level median property values as a weighted average of median property values within each of PUMA intersecting the county, where weights represent the percent of properties in each county estimated to come from a given PUMA.

Table 1: Years for which we have collected returns data for each NRI land use.

| | Land uses | | | | | |
Year	Crop	Pasture	CRP	Forest	Range	Urban
1982	X					
1987	X		X			
1992	X		X			
1997	X	X	X		X	
2002	X	X	X	X	X	X*
2007	X	X	X	X	X	X
2012	X	X	X	X	X	X
2015	X	X	X	X		X

We then merge county-level median property values with regional data from the US Census Survey of Construction. The Survey of Construction provides annual data on the number and characteristics of new housing units built in the US, by region. From the Survey of Construction, we collect regional data on average home sale price, lot value, and total lot size. We use lot value and average home sale price to construct region-level ratio of land value to total property value. We multiply county-level median property values by this ratio to obtain an estimate of median land values on newly developed properties. We then multiply this value by average lot size from the Survey of Construction to obtain estimates of median land value per acre on newly developed properties.

References

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.

Mihiar, Chris and David Lewis, "Climate, adaptation, and the value of forestland: A national Ricardian analysis of the United States," 2017.

_ **and** _ , "An Econometric Analysis of the Impact of Climate Change on Broad Land-Use Change in the Conterminous United States," 2019.