

B Sources and Construction of Variables

In this appendix chapter, I explain the sources and construction of all variables used in this paper. The main county-level data come from U.S. Census of Agriculture and Census of Population and posted at the website of ICPSR (Inter-university Consortium for Political and Social Research). Precipitation and temperature data come from Willmott and Matsuura (2001) at the University of Delaware, and are downloaded NOAA's website. Several additional variables borrowed from Hornbeck (2012) and Hornbeck and Keskin (2014) are posted at the authors' website. The program variations of the Great Plains Shelterbelt Project and the Great Plains Conservation Program are digitized by myself from Droze (1977) and documentations from USDA Soil Conservation Service. Hence, all the data used in this paper are publicly available, and all the raw data and my do files are available upon request.

Meanwhile, note that all analysis of this paper consistently uses county boundaries in 1910, so a few counties with boundary changes had minor adjustment in all variables. Assuming each variable is evenly distributed within these counties' boundaries in more recent years, spatial averages are calculated based on their 1910 boundaries, and are used instead of the raw numbers for years after the boundary changes.

B.1 Main County-level Data from ICPSR

The following set of variables come from ICPSR data sets, "Historical, Demographic, Economic, and Social Data: US, 1790-2000" (Haines, 2005) and "Great Plains Population and Environmental Data: Agricultural Data" (Gutmann, 2005), both collected from the U.S. Census of Agriculture and Census of Population. Variable construction mostly follows Hornbeck (2012) and Hornbeck and Keskin (2014):

Outcome Variables	Definition & Construction
Cropland / (Cropland + Pasture)	"Cropland" includes total cropland acreage harvested, fallowed, and failed; "Pasture" includes regular pasture, as well as cropland and woodland pastured. Available 1930-1992, although "Pasture" is missing and linearly interpolated for 1969.
log(Revenue from Crops / Farmland)	"Revenue from Crops" is directly reported or calculated from "total revenue \times crop fraction"; "Farmland" is directly reported or calculated from the summation of farmland acreage owned, managed, and tenant. Available 1910-1992.
log(Revenue from Animal Products / Farmland)	"Revenue from Animal Products" is directly reported, or calculated from the summation of revenue from livestock, dairy, and poultry, or calculated from "total revenue \times livestock fraction"; "Farmland" is the same as above. Available 1910-1992.
log(Total Revenue / Farmland)	"Total Revenue" is directly reported, or calculated from the summation of revenue from crops and from animal products; "Farmland" is the same as above. Available 1910-1992.
log(Yield of Wheat / Acre)	The yield and acreage of wheat is directly reported. Available 1910-1992.
log(Yield of Corn / Acre)	The yield and acreage of corn is the summation of corn for grain and corn for silage. Available 1910-1992.
log(Yield of Hays / Acre)	The yield and acreage of all hays is directly reported. Available 1910-1992.

Outcome Variables (cont.)	Definition & Construction
$\log(\text{Irrigated Acreage})$	"Irrigated Acreage" is directly reported. Available 1935-1992.
$\log(\text{Revenue from Crops} / \text{Cropland})$	"Revenue from Crops" is directly reported or calculated from "total revenue \times crop fraction"; "Cropland" includes total cropland acreage harvested, fallowed, and failed. Available 1910-1992.
$\log(\# \text{ of Cattle} / \text{Acre})$	"# of Cattle" is directly reported, and is divided by county area in 1910. Available 1910-1992.
$\log(\# \text{ of Pigs} / \text{Acre})$	"# of Pigs" is directly reported, and is divided by county area in 1910. Available 1910-1992.
$\log(\# \text{ of Chicken} / \text{Acre})$	"# of Chicken" is directly reported, and is divided by county area in 1910. Available 1910-1992.
$\log(\text{Expenditure on Feed})$	"Expenditure on Feed" is directly reported. Available 1910-1992 except 1930-1959.
$\log(\text{Revenue from Animal Products} / \text{Pasture})$	"Revenue from Animal Products" is directly reported, or calculated from the summation of revenues from livestock, dairy, and poultry, or calculated from "total revenue \times livestock fraction"; "Pasture" includes regular pasture, as well as cropland and woodland pastured. Available 1910-1992, although "Pasture" is missing and interpolated for 1969.
$\log(\text{Rural Population Density})$	"Rural Population" is directly reported, and is divided by county area in 1910. Available 1910-1980 decennially.
$\log(\text{Value of Equipment per Acre})$	"Value of Equipment" is directly reported, and is divided by farmland area. Available 1910-1992 except 1950-1959, 1964, 1982, and 1987.
Farmland/County area	"Farmland" is directly reported or calculated from the summation of farmland acreage owned, managed, and tenant; "County area" is reported in 1910. Available 1910-1992.
Woodland/County area	"Woodland" is directly reported woodland area; "County area" is the same as above. Available 1910-1992.

Other Variables	Definition & Construction
Population per 1,000 Acre	"Population" is directly reported, and is divided by county area in 1910. Data from 1910-1930 used.
Fraction of rural population	"Rural population" is directly reported, and is divided by "Population". Data from 1910-1930 used.
Fraction of farming population	"Farming population" is directly reported, and is divided by "Population". Data from 1910-1930 used.
No. of farms/1,000 Acre	"No. of farms" is directly reported, and is divided by county area in 1910. Data from 1910-1930 used.
Avg Farm Size (in Acre)	"Avg Farm Size" is farmland area divided by "No. of farms". Data from 1910-1930 used.
Area of corn/Cropland	"Area of corn" is the summation of acreage of corn for grain and corn for silage; "Cropland" includes total cropland acreage harvested, fallowed, and failed. Data from 1910-1930 used.
Area of wheat/Cropland	"Area of wheat" is directly reported, and "Cropland" is the same as above. Data from 1910-1930 used.
Area of hay/Cropland	"Area of hay" is directly reported, and "Cropland" is the same as above. Data from 1910-1930 used.
Area of cotton/Cropland	"Area of cotton" is directly reported, and "Cropland" is the same as above. Data from 1910-1930 used.
Area of oat, barley, and rye/Cropland	"Area of oat, barley, and rye" is the summation of acreage of oat, barley, and rye; "Cropland" is the same as above. Data from 1910-1930 used.
Land in Strip Cropping (%)	"Land in Strip Cropping (%)" is directly reported, and is divided by farmland area. Available in 1959-1969.
Cropland in Cover Crops (%)	"Cropland in Cover Crops (%)" is directly reported, and is divided by farmland area. Available in 1969-1992.

Other Variables (cont.)	Definition & Construction
Cropland in Conservation Reserve Program (%)	"Cropland in Conservation Reserve Program (%)" is directly reported, and is divided by farmland area. Available in 1987 and 1992.
Cropland diverted under Acreage Adjustment Program (%)	"Cropland diverted under Acreage Adjustment Program (%)" is directly reported, and is divided by farmland area. Available in 1978-1987.
Land in Diverted Commodity Program (%)	"Land in Diverted Commodity Program (%)" is directly reported, and is divided by farmland area. Available in 1992.
Cropland in Summer Fallow (%)	"Cropland in Summer Fallow (%)" is directly reported, and is divided by farmland area. Available in 1959-1992.

B.2 Variations in the Great Plains Shelterbelt Project

Key program variations of the Great Plains Shelterbelt Project (GPSP) are digitized by myself from two maps in Droze (1977), originally provided by U.S. Forest Services. See

https://play.google.com/books/reader?id=9_cTAAAYAAJ&printsec=frontcover&output=reader&hl=en&pg=GBS.PA122-IA1 and https://play.google.com/books/reader?id=9_cTAAAYAAJ&printsec=frontcover&output=reader&hl=en&pg=GBS.PR26.

Variables	Definition & Construction
Proportion protected by with shelterbelts	Proportion of each county within the "Major planting area" under the GPSP, as reported by the US Forest Service, i.e. the black area in Figure 1.
Proportion in the 100-mile-wide belt	Proportion of each county within the 100-mile-wide shelterbelt zone under the GPSP proposed by the US Forest Service, i.e. the belt with dashed boundary in Figure 1.

B.3 Climate Data

Precipitation and temperature data come from Willmott and Matsuura (2001) at the University of Delaware, and are provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their website at <https://www.esrl.noaa.gov/psd/>.

Variables	Definition & Construction
Total precipitation in recent two years (mm)	The summation of 24 months of precipitation including the calendar year when each wave of census data was collected and the previous calendar year.
Avg temperature in recent two years ($^{\circ}C$)	The pooled average temperature over the 24-month period including the calendar year when each wave of census data was collected and the previous calendar year.

B.4 Soil erosion and the Ogallala Aquifer

These additional covariates are borrowed from Hornbeck (2012) and Hornbeck and Keskin (2014) that are posted on the authors' website.

Variables	Definition & Construction
Proportion medium-eroded	Proportion of each county within the medium-eroded region in the Dust Bowl constructed in Hornbeck (2012).
Proportion high-eroded	Proportion of each county within the high-eroded region in the Dust Bowl constructed in Hornbeck (2012).
Proportion above the Ogallala Aquifer	Proportion of each county above the Ogallala Aquifer constructed in Hornbeck and Keskin (2014), i.e. the light blue area in Appendix Figure A.3.

B.5 The Great Plains Conservation Program

The GPCP designation is digitized by myself from the Soil Conservation Service (1982).

Variables	Definition & Construction
GPCP Designation	Proportion of each county within the GPCP designated region, i.e. the gray area in Appendix Figure A.4.