July, 2015

Three Essays on Land Use, Land Management, and Land Values in the Agro-Ecosystem

Wendong Zhang, *Iowa State University of Science and Technology*

Three Essays on Land Use, Land Management, and Land Values in the Agro-Ecosystem

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Wendong Zhang, B.S., M.A.

Graduate Program in Agricultural, Environmental and Development Economics

The Ohio State University

2015

Dissertation Committee:

Elena G. Irwin, Advisor

Brian E. Roe

Sathya Gopalakrishnan

Abstract

Over the past few years, U.S. agriculture and farmers have experienced a myriad of macroeconomic and environmental changes that have profound implications for the well-being of farm households and the farm sector. An expanding biofuels market and growing export demand from China and India have led to rising agricultural commodity prices since mid-2000s. However, during the same time period, the residential housing market collapsed in 2007-2008 and resulted in the subsequent Great Recession, which could impose a downturn pressure on the farmland market. In addition, growing water quality problems due to excessive agricultural nutrient runoff have severely compromised many ecosystem services and have led to stronger calls for more effective nutrient management policies from both policymakers and the public. Economic analyses of farmer decisions in this constrained and evolving environment are critical to understand how these changes have impacted farmer welfare and trade-offs with ecosystem and other societal benefits. Using individual-level data on farmland parcels and farmers from Ohio and Lake Erie basin, my dissertation examines how the recent residential housing market bust, expanding ethanol production, and rising environmental concerns over nutrient management have impacted farmers' land use, land management, and land transaction decisions and the implications of these changes for farmer welfare.

Farm real estate represents over 80% of the balance sheet of the farm sector and is the single largest item in a typical farmer's investment portfolio, and thus changes in farmland values could affect the welfare of farm households and the farm sector in general. The first two chapters examine the trends and determinants of farmland values in the Midwest in the 2000s decade. In particular, the first chapter identifies the impact of the recent residential housing market bust and subsequent economic recession on farmland values, using parcel-level farmland sales data from 2001-2010 for a 50-county region under urbanization pressure in western Ohio. My estimates from hedonic regressions reveal that farmland was not immune to the residential housing bust; the portion of farmland value attributable to urban demands for developable land was almost cut in half shortly after the housing market bust in 2009-2010. This chapter offers the first analysis of the magnitude of the structural break in the effect of urban influence on surrounding farmland values due to the recent housing market bust.

The second chapter investigates the capitalization of expanding biofuels market in surrounding farmland values. In particular, it tests for structural change in the relative effects of proximity to agricultural market channels before and after the construction of seven ethanol plants in or near western Ohio in late 2006 – early 2007. Instrumental variables regression on the matched sample demonstrates the positive capitalization of newly constructed ethanol plants. To the best of my knowledge, this chapter is the first to provide formal evidence of the effects of ethanol market expansion on farmland values during a strong recessionary time that exerted substantial downward pressure.

The last chapter examines the interplay between agriculture and the environment, as well as the trade-off between farmer welfare and benefits of ecosystem services resulting from alternative agri-environmental policies. Excessive agricultural nutrient runoff has severely compromised the sustainability of Lake Erie agri-ecosystem, however, current voluntary conservation payments policy have been proven insufficient for nutrient reduction. Using individual level data on farm, field, and farmer characteristics, the third chapter develops a structural econometric model of farmers' profit-maximizing output supply and input demand decisions, and quantifies the social welfare impacts of alternative nutrient management policies, including uniform and targeted fertilizer taxes. Results reveal that neither a fertilizer tax nor an education campaign could alone achieve the policy goal of 40% reduction in nutrient runoff into Lake Erie, although a uniform 50% fertilizer tax could lead to a 24% reduction in mean phosphorus application rates.. I also find that spatial targeting, such as phosphorus tax targeted towards ecologically sensitive subbasins, improves the cost-effectiveness of agri-environmental policies when only costs to farmers are considered; while a simpler policy such as a 50% uniform phosphorus tax would outperform other alternatives when the cost-effectiveness is measured as phosphorus reduction given net policy costs from an overall social welfare perspective.

Dedication

Dedicated to my beloved grandparents, my parents, my wife and daughter

for all the love, support, sacrifice and inspiration

Acknowledgments

This work is a collaborative effort that would not have been made possible without the support and assistance of many people.

First of all I would like to thank my wife Wei and my lovely daughter Lucy. They are the most wonderful blessings in my life and they have taught me how to be a better husband, a better father, and a better person through their love, laughter, sacrifice, and inspiration. I also want to thank my family for their unfailing support of me at each life stage, particularly my parents and parents in law. Words cannot describe what you have done for me. And a special thanks goes out in memory of my beloved grandparents, who are always at the bottom of my heart.

I want to thank Elena G. Irwin, my mentor and advisor, for always providing her unwavering research advice and support, for always making time for me even amid most busy days, for always looking out for opportunities for me to network, intern and write grants, for knowing when to push and when to listen, for navigating and guiding me through the stressful job market, for broadening my horizon as a researcher and a teacher. Looking back at the graduate school years, I deeply appreciate all the care, support and guidance Elena offered and I cannot think of a better Ph.D. advisor.

I would next like to thank my research support team, Brian Roe, Cindy J. Nickerson at ERS, Sathya Gopolakrishnan, Brent Sohngen, Jay Martin, Robyn Wilson, Abdoul Sam, Mark Partridge, H. Allen Klaiber, and Alan Randall for the feedback, patience, and expertise that you have provided for me to navigate my graduate school. I would also like to thank all of my graduate school peers and friends, especially Doug Wrenn, Matt Gnagey, Nic Irwin, Greg Howard, Xiaohui Tian, Minyu Zhou, and Michael Farren and many others who each contributed to my graduate school experience. Lastly, I would like to thank Ryan Williams and Vince Breneman of USDA ERS for support with the GIS data and variable generation for the first two chapters.

This research was gratefully supported by U.S. Department of Agriculture's Economic Research Service under cooperative agreement 58-6000-8-0065, NSF Coupled Human and Natural Systems grant (GRT00022685), as well as NOAA/Ohio Sea Grant.

Vita

June 2005 ....Shenxian No.1 High School, Shandong Province, China

July 2009..B.S. Environmental Science, Fudan University, China

June 2012 ....M.A. Economics, The Ohio State University

2012 to present ....Graduate Research Associate, Department of Agricultural, Environmental and Development Economics, The Ohio State University

Publications

Nickerson, C.J., and W. Zhang. 2014. "Modeling the Determinants of Farmland Values in the U.S." In J.M. Duke and J. Wu, ed. *The Oxford Handbook of Land Economics*. Oxford University Press, pp. 111-139.

Fields of Study

Major Field: Agricultural, Environmental and Development Economics

Table of Contents

List of Tables

List of Figures

Chapter 1: The Housing Market Bust and Farmland Values: Identifying the Changing Influence of Proximity to Urban Centers

Introduction

The recent residential housing market bust and subsequent economic recession have led to a dramatic decline in urban land values and housing values across the U.S. According to Standard & Poor's Case-Shiller repeat sales price index, residential property values in major metropolitan areas have declined by approximately 40% between 2007 and the end of 2008. Although farmland near urban areas provides a supply of land that could be developed for residential or commercial uses, a corresponding dip was not evident in farmland prices. Survey data reveals that farm real estate values witnessed a modest increase rather than a decline in many states over 2007 – 2009, including several with significant amounts of farmland subject to urban influence (Nickerson, et al. 2012). Favorable changes in factors that positively influence farmland values – including historically low interest rates that increase the attractiveness of farmland as an investment, and increasing demands for commodities (Gloy, et al. 2011; Schnitkey and Sherrick 2011; Wallander, et al. 2011)– may be masking declines attributable to changes in residential housing markets. These recent changes in urban housing values and the seeming immunity of nearby farmland values raise questions about the relationship between urban and farmland markets: what was the magnitude, if any, of the drag imposed by the urban residential housing market downturn on surrounding farmland values? Understanding how farmland values respond to fluctuations in competing land markets is of perennial policy interest, as changes in farmland values can affect the health of the farm sector and of farm household wellbeing. Farmland values represent over 80 percent of the value of farm sector assets, and farmland represents the largest asset in the typical farm household investment portfolio (Nickerson, et al. 2012).

Farmland in close proximity to urban areas typically sells for a premium relative to farmland farther away from urban areas - as demand for developable land induces developers to bid above the agricultural production value of land closest to urban areas (Capozza and Helsley 1989). Many empirical studies have shown that in more urbanized areas the demand for developable land for residential or commercial uses is the most significant nonfarm factor affecting farmland values (Cavailhès and Wavresky 2003; Hardie, et al. 2001; Livanis, et al. 2006; Shi, et al. 1997). However, most of these studies use aggregate county level data, which generates a very coarse representation of the spatial extent and magnitude of urban influence, and masks important differences in the influence of spatially disaggregate locational attributes on agricultural land values, such as parcel specific variation in distance to nearby city centers as a proxy for future development pressure. One exception is the study by Guiling, et al. (2009). They estimate a model that incorporates both county-level data and parcel characteristics, and find that urban influence on agricultural land values extended between 20 and 50 miles away from the closest urban centers, depending on the population and real income of the urban area. While Guiling, et al. (2009) demonstrated the spatial heterogeneity of urban influences in farmland markets, their model did not address the potential for substantial variation at a subcounty level (Bajari, et al. 2012), as well as the possibility of influences from multiple urban centers (Shi, et al. 1997).

The recent housing market boom-bust has sparked renewed interest in the impacts on land and house prices within and across metropolitan areas (Cohen, et al. 2012; Kuminoff and Pope 2013). Yet these studies on the influence of the housing boom and bust are limited to residential land and structure values, with no explicit representation of the impact on surrounding farmland that could be developed. A few recent farmland value studies have examined how changes in other non-land markets, such as demand for biofuels as an energy source, have affected farmland values but they did not consider the impact of changes in competing land markets (Blomendahl, et al. 2011; Henderson and Gloy 2009).

The aim of this study is to identify, at the parcel level, the total dollar value of proximity to urban centers (the "urban premium") and test for a structural change in these effects before and after the urban housing market bust that spanned from early 2007 through late 2008. I hypothesize that the urban housing market bust imposed significant downward pressure on urban demands for developable land and hence the urban premium that accrues to farmland near urban areas. This study uses spatially explicit parcel-level data on arms-length agricultural land sales from 2001 to 2010, a period which encompasses the housing market bust, for a 50-county region of western Ohio - almost all of which is subject to some degree of urban influence. This unique and spatially disaggregate dataset allows me to parse the data into pre (2000-2006) and post (2009-2010) time periods, and investigate the structural change in the effects of urban proximity on surrounding farmland values, yielding new insights into the impacts of changes in competing land markets on farmland values.

The parcel-specific urban premium metric explicitly considers the possibility of influences from multiple urban centers by adding three additional parcel-level measures of urban influences to the traditional metric "distance to nearest city", including surrounding urban population, the incremental distance to the second nearest city and a gravity index based on the nearest three cities to quantify the effects of multiple urban centers (Shi, et al. 1997).I also address the potential omitted variable bias embedded in the standard hedonic pricing approach by incorporating census tract fixed effects, which control for time-invariant unobserved spatial characteristics that could vary within a county and greatly affect the future development potential of farmland parcels, such as access to commuting opportunities, school quality, and air quality (Kuminoff and Pope 2013).

The main result provides evidence that the value of being within close proximity to urban centers on surrounding farmland values declined by an estimated 50 percent or so due to the recent residential housing market bust. On average, the urban premium for parcels under urban influence relative to a hypothetical parcel not subject to urban influence fell from $1,947 per acre before 2007 to $1,026 per acre shortly after the housing market bust, a decline of more than 40% to roughly 20% of per-acre farmland prices (without structures), respectively. The decline in the value of an urban premium due to the housing market bust was greater for parcels in closest proximity to cities. In addition, the results illustrate the importance of incorporating parcel level measures of the influences from multiple urban centers. The average parcel-level urban premium would be underestimated by as much as 17 percent before 2007 if measures accounting for multiple urban centers are omitted– suggesting multiple urban centers represent a significant portion of the urban premium at least in periods of strong housing market growth. Overall, this study makes at least two contributions to the literature on farmland valuation. First, to my knowledge, this study offers the first analysis of the magnitude of the structural break in the effect of urban influence on surrounding farmland values due to the recent housing market bust. In addition, this study develops a parcel-level measure of urban premium that explicitly accounts for the influences of multiple urban centers and shows that not accounting for the effects of multiple urban centers can result in a substantial undervaluation of the urban premium.

Conceptual Framework

Among the most influential theories that help explain the value of land is Ricardo's economic theory of rent (Ricardo 1996). Ricardo's key insight was that land which differs in quality and which is limited in supply generates rents that arise from the productive differences in land quality or in differences in location. The valuation of farmland subject to urban influence dates back to a model developed by Von Thünen in 1826, which posits that rent differentials for farmland also arise both from the value of commodities produced and the distance from central markets. In this model the

Ricardian rent is a decreasing function of the distance to the urban center, and land closer to the urban center earns higher rents because of reduced transportation costs. Farmland value is comprised of the net present value of economic returns to land. The model is written as

$$V_{it} = E_t \sum_s \frac{R_{is}}{(1+\delta)^{s-t}}, where\ s = t, t+1, \ldots \qquad (1)$$

In this formulation, the value of agricultural land parcel *i* at time *t* V_{it} is defined as the expected annual returns to farmland R discounted at rate δ. In many regions, farmland can earn returns not just from agricultural production and government payments, but also from "non-farm" sources such as wildlife viewing, hunting, and fishing. Principal among the non-farm sources of returns for farmland in close proximity to urban areas is the expected future rent increases arising from expected returns from future development for residential or commercial uses (Hardie, et al. 2001). Capozza and Helsley's (1989) seminal work laid the theoretical foundation for this literature and showed how the value of expected future rent increases could be quite large, especially near rapidly growing cities.

The study region - Western Ohio - is fairly homogenous in climatic conditions and opportunities for fishing or hunting opportunities, and hence little variation in generating recreational income is expected among the parcels. The area faces significant development pressure however, so I focus on returns arising from the option value of future land conversion from agricultural use to urban uses. Following Capozza and Helsley (1989), the value of an agricultural parcel *i* at time *t* under urban influence can be defined as

$$V_i(t) = \sum_{s=0}^{t^*} \frac{R_A(A_i,s)}{(1+\delta)^{s-t}} + \sum_{s=t^*}^{\infty} \frac{R_U(U_i,s)}{(1+\delta)^{s-t}}, \tag{2}$$

where t^* is the optimal timing of land use conversion from agricultural use to residential or commercial uses, R_A is the agricultural land rent, and R_U is the urban land rent net of the conversion costs. The first term represents the present value of agricultural rents up to t^*, which depends on the parcel-specific variables affecting agricultural productivity $\boldsymbol{A_{it}}$ such as soil quality, slope of the parcel, and proximity to agricultural market channels such as ethanol plants and grain elevators. The second term captures the present value of returns to urban development from the optimal conversion time onward, which depends on the location-specific urban influences variables $\boldsymbol{U_{it}}$ such as proximity to nearby cities, surrounding urban population, size of nearby multiple urban centers, and access to highway ramps and railway stations[1]. The recent decline in urban housing market demands may greatly diminish the urban option conversion value of agricultural land relative to the preceding period of high housing demand, and as a result, a declining significance of the urban influence variables $\boldsymbol{U_{it}}$ in shaping surrounding farmland values is expected between the two periods.

Econometric Procedures

The Hedonic Price Method

Hedonic models are a revealed preference method based on the notion that the price of a good or parcel in the marketplace is a function of its attributes and characteristics. With

[1] The increased access to customers could also influence farmland values by increasing expected agricultural returns. However this effect may be most relevant when there are many dairy, fruit and vegetable farms, which is not the case for my study region.

Rosen's (1974) seminal work as a backdrop (Rosen 1974), the hedonic price method has become the workhorse model in the studies of real estate or land values (Palmquist 1989), and the determinants of farmland values. Numerous applications of hedonic models applied to farmland markets have examined the marginal value of both farm and non-farm characteristics of farmland, including soil erodibility (Palmquist and Danielson 1989), urban proximity (Shi, et al. 1997), wildlife recreational opportunities (Henderson and Moore 2006), zoning (Chicoine 1981), and farmland protection easements (Nickerson and Lynch 2001). The farmland returns R_{it} in equation (2) can be approximated by a linear combination of parcel attributes and location characteristics using Taylor expansion. Hedonic models are commonly specified in log-linear form[2], which is defined as

$$\log(V_{it}) = \beta_0 + \beta_A{}'\boldsymbol{A_{it}} + \beta_U{}'\boldsymbol{U_{it}} + \tau_t + \varepsilon_{it}, \qquad (3)$$

where τ_t is time fixed effects which captures the temporal variations in returns and discount factor, and ε_{it} is the remaining normally distributed error term, and the agricultural land values $V_{it}$are approximated by the nominal sale prices per acre of the agricultural land without structures.

[2] I choose a log-linear functional form rather than the Box-Cox transformation of both dependent and independent variables because my interaction terms of urban influence have many zeros: Box-Cox transformation requires positive values. A robustness check using a Box-Cox transformation of the dependent variable (sale prices of farmland parcels) only yields a Box-Cox transformation parameter of 0.27, which is close to 0 as the parameter implied by log-linear functional form; also, the Box-Cox regression yields qualitatively similar results. I also add one robustness check using log-log specification and the results shown in Table 6 column (d) yield similar conclusions.

In this hedonic setting, agricultural land is regarded as a differentiated product with a bundle of agricultural quality and location characteristics, and each characteristic is valued by its implicit price.

Incorporating the Hedonic Model with Localized Spatial Fixed Effects

Despite its popularity, the hedonic pricing method suffers from a number of well-known econometric problems. Foremost among them, the researcher cannot directly observe all land characteristics that are relevant to farmers and developers, and omitted variables may lead to biased estimates of the implicit prices of the observed attributes (Bajari, et al. 2012). In the case of agricultural land under urbanization pressures, access to employment opportunities, school quality, and air quality could greatly affect future development potential and could vary significantly within a county, but be difficult to measure (Kuminoff and Pope 2013). For agricultural land parcels under no immediate urban conversion pressures, some other significant unobserved characteristics may also exist, such as access to public services and local climatic conditions. These characteristics are relatively homogenous within a census tract, so I address the omitted variable bias problem by incorporating local-level spatial fixed effects at the census tract level, which are denoted as θ_j (where the subscript *j* represents the census tract):

$$\log(V_{it}) = \beta_0 + \beta_A{}'\boldsymbol{A_{it}} + \beta_U{}'\boldsymbol{U_{it}} + \tau_t + \theta_j + \varepsilon_{it}, \qquad (4)$$

Previous studies have shown that coarser fixed effects at the county level may exclude too much intra-county variation and thus perform poorly in controlling for unobserved spatial heterogeneity (Anderson and West 2006). The localized spatial fixed effects I use here at the census tract level have been shown to effectively remove most of the time-invariant omitted variable bias, such as spatial autocorrelation (Abbott and Klaiber 2011). In addition, regression diagnostic techniques (e.g. *Moran's I* and *Geary's C*) are used as robustness checks to test for spatial autocorrelation in the residuals.

Construction of the Urban Premium

To better quantify the structural break in the effect of urban influences on surrounding farmland values induced by the housing market bust, I develop a parcel level measure of an "urban premium". This metric quantifies for each parcel, relative to a hypothetical agricultural land parcel with no urban influence, the total dollar value resulting from being located closer to urban areas. This urban premium measure consists of four distinct parts: value derived from being closer to the nearest city with at least 40,000 people[3] than the reference parcel, additional value derived from being within proximity to multiple urban centers – including incremental distance to the second nearest city, the positive effects resulting from surrounding urban population within 25 miles of the parcel centroid, and the value derived from total weighted population of the three nearest cities captured in a gravity population index. With these measures, I are able to identify the parcel-level structural change in the influence of urban premium before and after the

[3] In this study, I define cities as those with at least 40,000 people, and this threshold is used throughout the paper for distance calculations unless noted otherwise. While 50,000 people are used by the U.S. Census Bureau to define urbanized areas, I choose the threshold of 40,000 people because some core cities in Ohio Metropolitan Statistical Area such as Lima, OH have less than 50,000 people. The results are similar when a 50,000 threshold is used.

housing market bust. To construct this metric, the coefficients from the hedonic model with spatial fixed effects are used:

$$\log(V_{it}) = \beta_0 + \beta_A{}' \boldsymbol{A_{it}} + \beta_{U\_boom}{}' \boldsymbol{U_{it}} + \beta_{U\_bust}{}' \boldsymbol{U_{it}} * D_{t\_bust} + \tau_t + \theta_j + \varepsilon_{it}, \quad (5)$$

where D_{t_bust} is a binary time dummy indicating that the parcel is sold after the housing market bust. My main specification uses 2001 to 2006 as the pre (boom) period, and 2009 to 2010 as the post (bust) period. The pre- and post- periods were determined based on changes in the residential housing price indexes in Cleveland and Cincinnati metropolitan areas. These indexes exhibited rapid declines through the end of 2008, and a relative leveling off in 2009 and 2010 (Lincoln Institute of Land Policy 2012). The years 2007 and 2008 are treated as a transition period.

The parcel level urban premium is calculated as the difference between the predicted prices $\exp\left(\widehat{\log(P_{it})} + \widehat{\sigma_\epsilon^2}/2\right)$ using actual distance and population variables $\boldsymbol{U_{it}}$ for one parcel and the predicted prices $\exp\left(\dddot{\log(P_{it})} + \widehat{\sigma_\epsilon^2}/2\right)$ using distance and population variables $\boldsymbol{\bar{U}}$ of the reference parcel with no urban influence, where $\widehat{\sigma_\epsilon^2}$ is the corresponding mean squared error (MSE) from the regression model following equation (5):

$$\widehat{\log(P_{it})} = \widehat{\beta_0} + \widehat{\beta_A}{}' \boldsymbol{A_{it}} + \widehat{\beta_{U_{boom}}}{}' \boldsymbol{U_{it}} + \widehat{\beta_{U_{bust}}}{}' \boldsymbol{U_{it}} * D_{t_{bust}} + \widehat{\tau_t} + \widehat{\theta_j} \quad (6)$$

$$\dddot{\log(P_{it})} = \widehat{\beta_0} + \widehat{\beta_A}{}' \boldsymbol{A_{it}} + \widehat{\beta_{U\_boom}}{}' \boldsymbol{\bar{U}} + \widehat{\beta_{U\_bust}}{}' \boldsymbol{\bar{U}} * D_{t\_bust} + \widehat{\tau_t} + \widehat{\theta_j} \quad (7)$$

$$urban\ premium = \exp\left(\log(P_{it}\widehat{(\boldsymbol{U_{it}})}) + \widehat{\sigma_\epsilon^2}/2\right) - \exp\left(\log(\dddot{P_{it}}(\boldsymbol{\bar{U}})) + \widehat{\sigma_\epsilon^2}/2\right) \quad (8)$$

Guiling, et al. (2009) estimated the extent of urban influence using parcel level data in Oklahoma, and found that for a city with around 50,000 residents, the urban influence on farmland prices extends 45 miles from the city center. Semiparametric regressions using my data in Ohio reveal that the effects of urban influence become negligible around 60 miles away from the nearest city center, and the effects of the *incremental distance to the second nearest city center*[4] are no longer evident beyond 40 miles[5]. As a result, the distance and population variables for the reference parcel in this study are 60 miles for *the distance to nearest city*, 40 miles for *the incremental distance to the second nearest city*, and zero for surrounding urban population and gravity index. Using this definition, my measure of the urban premium is constructed relative to the hypothetical, rural parcel whose urban influence variables are denoted as $\bar{\mathbf{U}}$[6]. In my study region of Ohio, this metric is always positive for all the agricultural parcels.

[4] The incremental distance to second nearest city is defined as the difference between the distance from the second nearest city center and the distance from the nearest city center. For example, a parcel located 10 miles away from the nearest city center and 30 miles away from the second nearest city center will have an incremental distance to the second nearest city of 20 miles.

[5] The semiparametric regressions are estimated using the semip() function from the McSpatial package in R, and the model specification is following equation (4) with county fixed effects, with either distance to nearest city center or incremental distance to the second nearest city center estimated nonparametrically using locally weighted regressions. A robustness check using 50 miles and 30 miles for the thresholds of distance to nearest city center and incremental distance to second nearest city center respectively yield qualitatively similar results regarding the parcel-level urban premium.

[6] Numerically $\bar{\mathrm{U}}$ for this hypothetical parcel is assumed to be 60 miles away from nearest city center, 40 additional miles from the second city center, and 0 for surrounding urban population and the gravity index.

Data

Western Ohio hosts the vast majority of the state's agricultural land and provides an excellent laboratory to study the structural change in the determinants of farmland values that was precipitated by the residential housing bust. Ohio was hit hard in the housing market bust and accompanying recession, as evidenced by the sharp decline in residential housing prices for its metropolitan areas in 2007 and 2008 (Lincoln Institute of Land Policy 2012). To analyze the impact of the housing market bust, I assembled a detailed database of 21,342 arm's length agricultural land sale records for 50 western Ohio counties obtained from county assessors' offices and from a private data vendor.

The sample was further screened to eliminate farmland parcels under no or little urban influences: parcels were dropped if they were both outside the Core Based Statistical Area counties[7] and more than 10 miles away from the edge of the nearest city (with a population at least 40,000 people). In addition, only those agricultural parcels sold at arm's length between 2001 and 2010 were retained. These agricultural parcel sale records were merged with georeferenced parcel boundaries, or were geocoded based on property addresses using ArcGIS when georeferenced parcel boundaries were not available[8]. In the

[7] Core Based Statistical Areas (CBSAs) are defined by the U.S. Census Bureau as "consist[ing] of the county or counties or equivalent entities associated with at least one core (urbanized area or urban cluster) of at least 10,000 population, plus adjacent counties having a high degree of social and economic integration with the core as measured through commuting ties with the counties associated with the core. The general concept of a CBSA is that of a core area containing a substantial population nucleus, together with adjacent communities having a high degree of economic and social integration with that core."

[8] For these geocoded parcels, the parcel boundaries are proxied by square-shaped parcels with the same acreage.

hedonic regressions, parcels that sold between 2001 and 2006 were treated as sold during the pre (boom) period, and in the post (bust) period if sold in 2009-2010.

Construction of the dependent variable is a common problem in farmland value studies, given that sale prices reflect the value of both land and buildings including farm structures, residential dwellings, or both (Nickerson and Zhang 2014). Because I do not have data on the quantity and quality of buildings, I constructed a sales price for farmland only to use as the dependent variable. Similar to Guiling, et al. (2009) who subtracted the value of buildings from farmland sales prices, I calculated the sales price for farmland only as the original sales price times the ratio of the percentage of *assessed values of land only* over *total assessed values of land and buildings*. This assumes the portion of sales price attributable to land only can be approximated based on the contribution of assessed value of land to the total assessed value of land plus buildings. Parcels were dropped when the estimated sales price for farmland only was above $20,000/acre or below $1,000/acre. Figure 1 shows a plot of the filtered sample consisting of 12, 432 valid parcel transactions. As is evident from the figure, these data are widely distributed over the entire region. The temporal trends of farmland prices with and without structures for these filtered parcels are plotted in Figure 2, and the drastic decline experienced in the residential housing markets is not evident. A modest decline in average farmland prices with structures (the farm real estate values) from the mid-2000s is noticeable. The average nominal farmland sale prices without structures stayed fairly

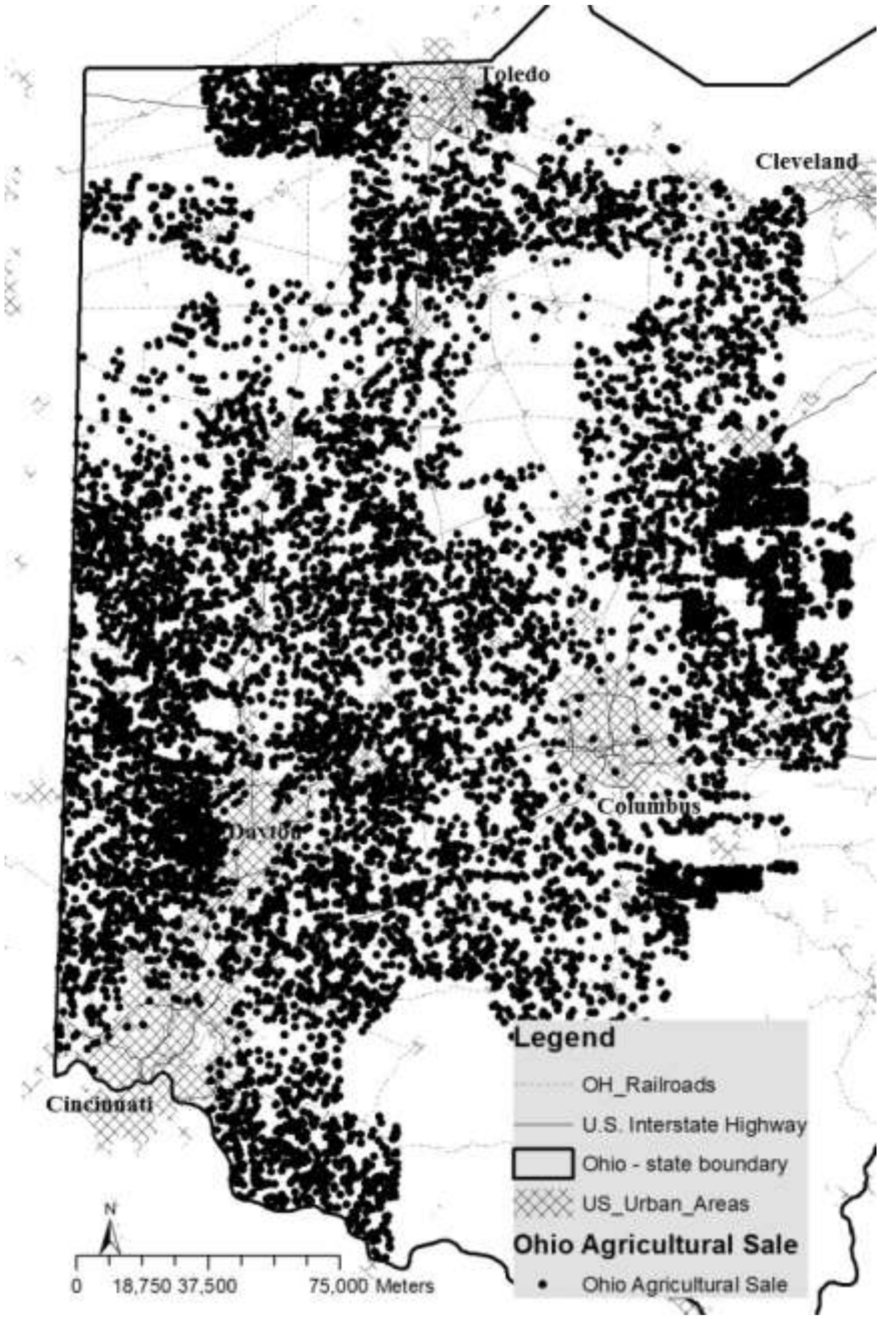


Figure 1. Farmland Land Sales under Urban Influence in Western Ohio 2001-2010

constant around $4,500 per acre over the 2000 decade, yet a noticeable dip occurred between 2008 and 2009.

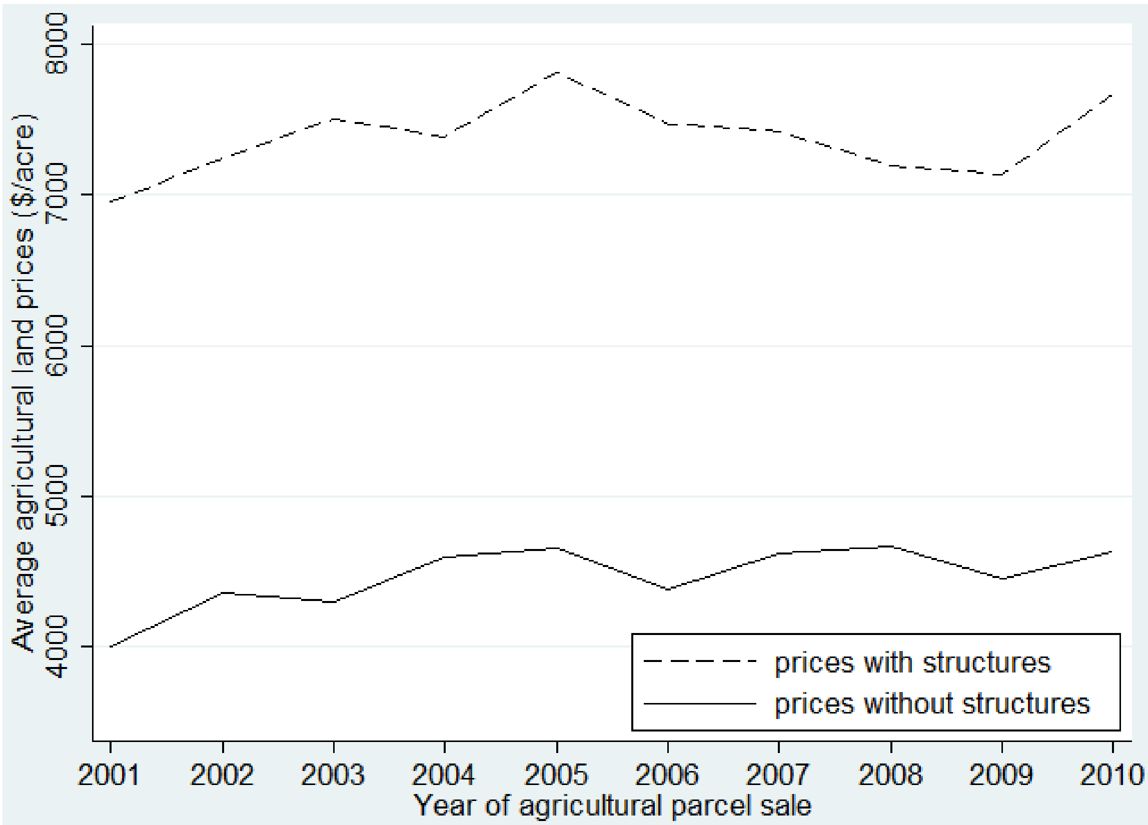


Figure 2. Distribution of Real Arms-length Farmland Prices 2001-2010 in Western Ohio

Data on parcel attributes and location characteristics were obtained largely from the U.S. Department of Agriculture Natural Resources Conservation Service's GeoSpatial Data Gateway (USDA GeoSpatial Data Gateway, 2012), including the Census TIGER/Line Streets, National Elevation Dataset, National Land Cover Dataset (NLCD), and Soil

Survey Spatial Data (SSURGO). Additional data on locations of cities and towns in Ohio were obtained from the Ohio Department of Transportation (2012). I also used Census Block Shapefiles with 2010 Census Population and Housing Unit Counts (U.S. Census TIGER/Line 2012) to calculate the surrounding urban population. Data on ethanol plants, grain elevators and agricultural terminal ports were obtained from the Ohio Ethanol Council (2012), the Farm Net Services (2012) and the Ohio Department of Agriculture (2012). Using these data and ArcGIS software, I were able to create the parcel attributes and location characteristics. Table 1 reports summary statistics for these variables. Several variables in Table 1 are self-explanatory; however, a number of explanations are in order. First, the variable *National Commodity Crops Productivity Index (NCCPI)* is an interpretation in the National Soil Information System (NASIS). Specifically, the interpretation is based on natural relationships of soil, landscape, and climate factors and assigns productivity ratings for dry-land commodity crops, where the most desirable properties, landscape features and climatic conditions lead to larger values of *NCCPI* (see Dobos, et al. (2008) for details). The *percentage of prime farmland* variable is based on the suitability of soils for most kinds of field crops: for each parcel, the percentage measure of land area in prime soil is calculated. The grain elevators and agricultural terminals were in operation before the start date of this study, and thus the distances to these two types of agricultural delivery points are constant over the study period. However, all of the six ethanol plants in Western Ohio did not start operations until 2008.

	Unit	Mean	Std. Dev.	Min.	Max.
General Parcel Attributes					
Sales price per acre (with structures)	Dollars	7374.65	6037.55	1106.2	31260.4
Sales price per acre (without structures)	Dollars	4456.96	3497.43	1000.16	19999.7
Assessed land value % of total assessed	%	72.87%	29.96%	5.38%	100.00%
Total acres	Acres	46.83	64.68	0.14	2381
Sale year	Year	2004.96	2.67	2001	2010
Agricultural Profitability Influence Variables					
National Commodity Crops Productivity Index	Number	5739.35	1571.55	0	8800.8
Cropland % of parcel	%	54.49%	37.80%	0.00%	100.00%
Prime soil % of parcel	%	37.52%	36.18%	0.00%	100.00%
Steep slope		0.42	0.71	0	3
Distance to nearest ethanol plant	Miles	29.65	13.89	0.55	69.84
Distance to nearest grain elevator	Miles	8.18	6.88	0.03	55.27
Distance to nearest other agricultural terminal	Miles	31.37	14.66	0.13	74.62
Forest area % of parcel	%	16.38%	26.84%	0.00%	100.00%
Wetland area % of parcel	%	0.34%	2.92%	0.00%	100.00%
Urban Influence Variables					
Distance to nearest city center with over 40,000 people	Miles	22.56	10.57	0.12	57.39
Distance to nearest city center * after 2008	Miles	7.36	12.37	0	55.13
Incremental distance to second nearest city with at least 40k people	Miles	15.10	13.72	0.01	63.59
Incremental distance to second city * dummy of sale after 2008	Miles	4.68	10.24	0	63.57
Total urban population within 25 miles	Thousands	312.83	236.60	64.77	1187.38
Total urban population * after 2008	Thousands	89.24	176.58	0	1184.37
Gravity index of three nearest cities		1326.87	39204.4	62.14	4255332
Gravity index * after 2008		674.62	39194.53	0	4255332
Building area % of parcel	%	3.32%	12.45%	0.00%	100.00%
Distance to highway ramp	Miles	3.21	2.05	0	11.94
Distance to railway station	Miles	3.07	1.81	0.01	11.25
Number of observations			12432		

Table 1. Summary Statistics of Agricultural Land Sales under Urban Influences in Western Ohio

As a result, I assume the positive value of proximity to ethanol plants did not get capitalized before 2007 and thus the variable *distance to nearest ethanol plant* is interacted with a post 2008 time dummy.

Several measures of urban influences are considered: *distance to nearest city center* captures the importance of urbanized areas as a commuting hub or sources of non-farm income, and the potential for future urban development. *Surrounding urban population within 25 mile-radius for each parcel* also represents nearby demand for future land conversion to urban uses. The *incremental distance to second nearest city* is a measure commonly used in housing and labor market studies on Central Place Theory and urban hierarchy to capture the additional value of influences from multiple urban centers (Partridge, et al. 2008). The *incremental distance to second nearest city* (see footnote *iv*), the *surrounding urban population*, and the *gravity index* account for the aggregate urban influences resulting from multiple urban centers. The *gravity index* is calculated as the weighted average of population divided by distance squared for the nearest three cities following Shi, et al. (1997). Together, these four measures capture the most salient aspects of urban influences and are used to construct the urban premium described in section III.c. Some additional measures related to urban influences are also considered as controls. *The percentage of building area within a parcel* is included to capture any unobserved value of farm structures and houses that may remain in my "land only" measure of sales price. The unobserved value captured by t*he percentage of building area within a parcel* is more closely tied to heterogeneous preferences of houses or agricultural production needs than to urban proximity, and thus is excluded in the construction of the urban premium. The *distance to the nearest highway on-ramp* and *the distance to the nearest railway station* represent the additional value of being in close proximity to the interstate network and railway system, respectively. Variables on proximity to road networks are relatively homogenous among parcels and across time in my study region; in addition, they are shown to have a minor impact compared to the four main urban influence variables described earlier in this paragraph. As a result, these two road network proximity variables are not used to construct the urban premium.

Results and Discussion

Table 2 presents the results of my tests for structural change in the effect of urban influence using a hedonic model with 505 census tract fixed effects, denoted as the default model – model 0. The key variables are the urban influence variables such as *distance to nearest city* and their interactions with the post-2008 dummy. The *post-2008 dummy* is defined to be 1 if the parcel is sold after 2008. The interaction terms include the four urban influence variables mentioned in section III.c. Compared to the effects before 2007, the coefficients of these interaction terms indicate the significance and the magnitude of the structural break in the effects of urban influence after the housing market bust. The *distance to nearest city center* is further decomposed into whether the

Model	Model 0	
	Coef.	Std. Err.
Intercept	8.0343***	0.1743
Assessed land value % of total assessed	0.4270***	0.0226
Total acres	-0.0054***	0.0002
Total acres squared	2.95E-06***	1.26E-07
Agricultural Profitability Influence Variables		
National Commodity Crops Productivity Index	1.27E-05**	5.16E-06
Prime Soil area % of parcel	0.0473**	0.0206
Steep slope	-0.0112	0.0114
Forest area % of parcel	0.0053	0.0303
Wetland area % of parcel	-0.2851	0.2198
Distance to nearest ethanol plant * Post 2008 dummy	-0.0023*	0.0014
Distance to nearest grain elevator	-0.0011	0.0014
Distance to nearest other agricultural terminal	-0.0040***	0.0006
Urban Influence Variables		
Distance to city center*within 10 miles from urban boundary	-0.0088***	0.0013
Distance to city center*within 10 miles from urban boundary*Post 2008 dummy	0.0051**	0.0026
Distance to city center*beyond 10 miles from urban boundary	-0.0091***	0.0012
Distance to city center*beyond 10 miles from urban boundary*Post 2008 dummy	0.0057***	0.0025
Incremental distance to second nearest city center	-0.0035***	0.0008
Incremental distance to second nearest city center*Post 2008 dummy	0.0027*	0.0016
Total surrounding population within 25 miles	2.30E-04***	4.64E-05
Total surrounding population within 25 miles*Post 2008 dummy	9.57E-05	1.20E-04
Gravity index of three nearest cities	2.14E-05***	5.68E-06
Gravity index of three nearest cities*Post 2008 dummy	-2.20E-05***	5.71E-06
Building area % of parcel	0.1014**	0.0513
Distance to highway ramp	-0.0050	0.0033
Distance to railway station	-0.0003	0.0036
Year fixed effects	yes	
Census tract fixed effects	yes	
Adjusted R-square	0.2335	
Root mean squared error	0.6240	
Number of observations	10604	

Continued

Table 2. Hedonic Regression with Structural Changes in Urban Influence Variables

Table 2 continued
Note: the dependent variable in this model is the log of per-acre agricultural land prices without structures. *, **, and *** indicates the coefficient is significant at 10%, 5% and 1% level, respectively. 505 census tract fixed effects are included in the model.

parcel is within or beyond 10 miles from the boundary of an urbanized area with at least 40,000 people[9]. This term allows me to assess whether the marginal effect of distance to city is significantly different for parcels within 10 miles of the boundary of population centers, which previous research suggests is a point beyond which the effect of urban influences on farmland values is much less evident (Nickerson, et al. 2012).

Several points are notable regarding the urban influence variables and their effects. Before 2007, all of the coefficients of the four major urban influence variables are significant at the 1% level, confirming previous findings that urban influence is the most important non-farm factor in shaping farmland values in areas facing urbanization pressures. The biggest of these contributors is *the distance to nearest city center*, whose effect is almost twice as big as that of *incremental distance to second nearest city center*. The magnitude of the effect of distance before 2007 is a 0.88% increase in surrounding farmland values for each one-mile reduction in distance to nearest city center, and is comparable to the findings of previous studies (Ma and Swinton 2011). All else equal, the positive benefit per acre resulting from being closer to the nearest city declined from a

[9] The "within 10 miles" binary variable equals one for parcels inside or within 10 miles of the boundary of an urbanized area, and is zero otherwise. The "beyond 10 miles" binary variable equals one for parcels more than 10 miles of the boundary of an urbanized area, and is zero otherwise. As explained in footnote iii, I use 40,000 people as the threshold of urbanized areas, and similar results are found when a 50,000 or 25,000 threshold was used.

Model	Model I[#]	Model II	Model III	Model IV	Model V	Model VI	Model VII
Dist_City*within 10 miles	-0.0095***	-0.0103***		-0.0085***	-0.0119***	-0.1001***	-0.0096***
	(0.0013)	(0.0011)		(0.0014)	(0.0018)	(0.0013)	(0.0015)
Dist_City*within 10 miles	0.0048*	0.0047**		0.0052*	0.0045**	-0.0024	0.0004
*Post 2008 dummy	(0.0026)	(0.0022)		(0.0027)	(0.0025)	(0.0029)	(0.0017)
Dist_City*beyond 10 miles	-0.0090***	-0.0120***		-0.0089***	-0.0121***	-0.0100***	-0.0098***
	(0.0012)	(0.0008)		(0.0012)	(0.0018)	(0.0012)	(0.0013)
Dist_City*beyond 10 miles	0.0060**	0.0053***		0.0060**	0.0051**	-0.0033	0.0008
*Post 2008 dummy	(0.0025)	(0.0018)		(0.0026)	(0.0024)	(0.0026)	(0.0016)
Dist_City			-0.0091***				
			(0.0012)				
Dist_City*Post 2008 dummy			0.0055**				
			(0.0024)				
Incre Dist_2nd City	-0.0036*		-0.0035*	-0.0034*	-0.0072***	-0.0038***	-0.0041***
	(0.0008)		(0.0008)	(0.0008)	(0.0012)	(0.0008)	(0.0009)
Incre Dist_2nd City	0.0024		0.0027*	0.0033**	0.0022	-0.0004	-0.0010
*Post 2008 dummy	(0.0016)		(0.0016)	(0.0008)	(0.0016)	(0.0017)	(0.0011)
Urban popu within 25 miles	0.0002***		0.0002***	0.0003***	7.55E-06	0.0002***	0.0002***
	(4.69E-05)		(4.49E-05)	(4.83E-05)	(5.44E-05)	(4.63E-05)	(5.12E-05)
Urban popu within 25 miles	0.0001		8.23E-05	4.19E-05	0.0002	-0.0004***	-1.60E-05**
*Post 2008 dummy	(0.0001)		(0.0001)	(0.0001)	(0.0001)	(0.0001)	(7.26E-05)

Continued

Table 3. Robustness Checks of the Hedonic Regressions

Table 3 continued

Gravity index	2.09E-05***		2.12E-05***	1.78E-05***	2.06E-05***	1.95E-05***	1.87E-05***
	(5.68E-06)		(5.67E-06)	(5.92E-06)	(5.79E-06)	(5.68E-06)	(6.04E-06)
Gravity index*Post 2008 dummy	-2.10E-05***		-2 05	-1.90E-05***	-2.10E-05***	-1.89E-05***	-1.90E-05***
	(5.71E-06)		(5.70E-06)	(5.95E-06)	(5.82E-06)	(5.68E-06)	(6.05E-06)
Building area % of parcel	0.1001*	0.1266**	0.1015**	0.1386***	0.1009**	0.0657	0.0973**
	(0.0513)	(0.0511)	(0.0512)	(0.0534)	(0.0500)	(0.0535)	(0.0481)
Distance to highway ramp	-0.0055*	-0.0071**	-0.0051*	-0.0052	-0.0042	-0.0051	-0.0036
	(0.0033)	(0.0033)	(0.0033)	(0.0034)	(0.0032)	(0.0033)	(0.0031)
Distance to railway station	0.0005	0.0018	0.0004	0.0018	0.0023	0.0005	-4.42E-06
	(0.0036)	(0.0036)	(0.0036)	(0.0037)	(0.0035)	(0.0036)	(0.0034)
County fixed effects					Yes		
Census tract fixed effects	Yes	Yes	Yes	Yes		Yes	Yes
The post period is 2008 only						Yes	
Shifting the year of change to 2005							Yes
Root mean squared error	0.6239	0.6239	0.6239	0.6502	0.6169	0.6227	0.6203
Adjusted R-square	0.2336	0.2314	0.2336	0.5033	0.2508	0.2355	0.2197
Number of observations	10604	10604	10604	10604	10604	10350	11723

[#]: Model I distinguishes parcels not by within 10 miles of the boundaries of urbanized areas with at least 50,000 people, but by within 20 miles of the boundaries of urbanized areas with at least 100,000 people. Standard Errors are in parentheses. The dependent variable in this model is the log of per-acre agricultural land prices without structures. *, **, and *** indicates the coefficient is significant at 10%, 5% and 1% level, respectively. All models include year fixed effects.

significant effect of $30.92 per mile before 2007 to an insignificant $12.97 per mile effect after the housing market bust, an almost 60 percent reduction. In other words, due to the housing market bust, the single largest source of urban influence became insignificant in shaping surrounding farmland values, at least in the immediate short run. The decline is universal across parcels that are located within 10 miles from the boundary of urbanized areas or that are farther away. In addition, the effects of multiple urban centers are no longer significant after 2007[10]. In 2009 and 2010, the only urban influence variable that is still significant is the surrounding urban population.

The validity of the results is tested using multiple robustness checks shown in Table 3[11]. Different specifications and different samples are used to construct these robustness checks. Model I changes "within 10 miles from the boundary of urbanized areas with at least 50,000 people" to "within 20 miles from the boundary of urbanized areas with at least 100,000 people", because semiparametric analysis reveals that the effects of large urban centers (with at least 100,000 people) may not disappear until 20 miles away from its boundary[12]. I only include *the distance to nearest city center* in model II to investigate

[10] The significance of the urban influence variables after 2008 is tested using joint-restriction Wald test. For example, the F-statistic of *distance to nearest city center* + *distance to nearest city center * post 2008 dummy* reveals that the proximity to nearest city center is still significant at the 1% level after 2008, although the magnitude of the coefficient is reduced. However, similar results show that the other three urban influence variables, *incremental distance to second nearest cities, surrounding urban population,* and *gravity index,* are no longer significant after the housing market bust at the 10% level.

[11] Additional robustness checks using township fixed effects reveal almost identical results as the main specification shown in Table 2 and thus were not included in Table 3. These results are shown in Tables 6 and 7 column (b).

[12] See Figure 3 for the coefficient of distance to the boundary of urbanized areas from semiparametric regressions. Other regression results and corresponding figures for semiparametric regressions used to define the hypothetical parcel subject to no influence are available from the authors upon request.

the significance and contribution of the other three measures of multiple urban influences in the total urban premium; model III does not distinguish parcels within 10 or 20 miles from the boundary of urbanized areas from those beyond the cutoff; models IV uses the log of nominal farmland prices with structures as the dependent variable; model V uses county fixed effects rather than census tract fixed effects; model VI tests my assumption of the time lag effects by using parcels sold in 2008 as the post period group; and model VII assumes the housing market bust happened in 2005 rather than 2007-2008 to examine the possibility of falling urban influence due to factors other than the housing market bust, such as preference changes.

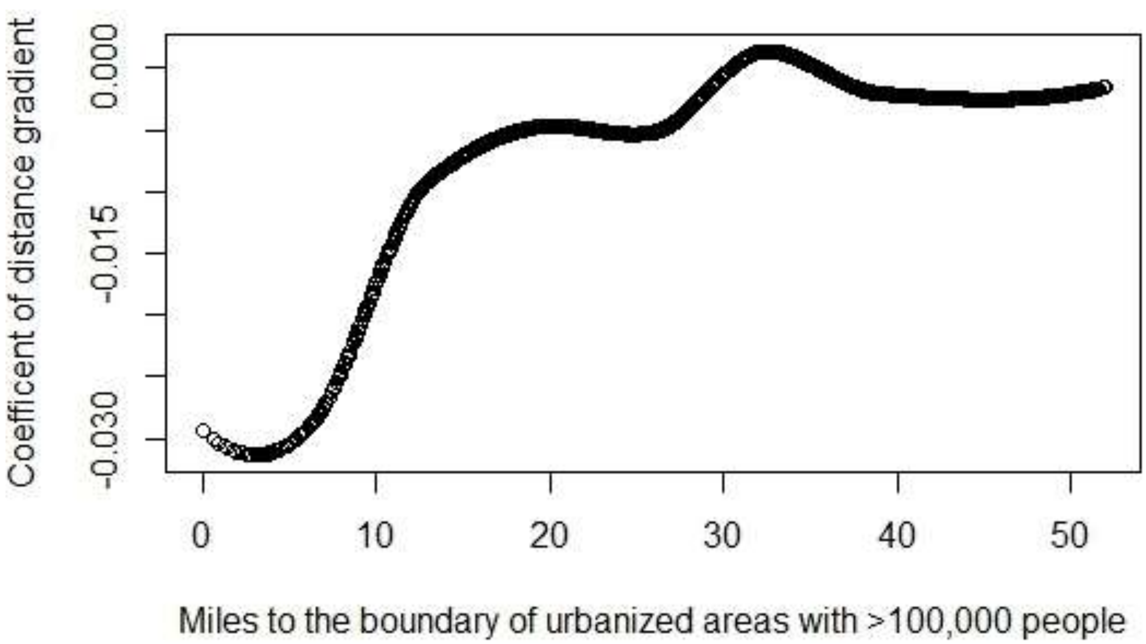


Figure 3. Semiparametric Analysis – Miles to the Boundary of Urbanized Areas with At Least 100,000 People

The results across different specifications can be grouped into four groups. First, models I, II, and III using census tract fixed effects and model IV using farmland prices with structures yield similar results as the main specification in Table 3: the impact of all of the four major urban influence variables except the surrounding urban population switched from significant before 2007 to negligible in 2009 and 2010. For example, model III reveals that the effect of proximity to the second nearest city center after the housing market bust (that is, the sum of the coefficients on *Incre Dist_2nd City* and *Incre Dist_2nd City*post-2008 dummy*) is statistically insignificant[13]. Secondly, in model V with county fixed effects, the proximity variables to nearest and second nearest city center are both significant throughout the decade, however the evidence of structural change is consistent: the effects are greatly reduced after the housing market bust. Comparisons of model V and others also show that county fixed effects obscured the value of some important urban influence variable, namely *surrounding urban population* even before the housing market bust. In addition, in model V with county fixed effects, the magnitude of the coefficient on *distance to the nearest city center* is about 30 percent higher than that in other model specifications with census tract fixed effects – both before and after 2007, suggesting a higher estimate of the urban premium in models with county fixed effects. This higher estimate could result from omitted characteristics at the subcounty level; however, it may also be possible that due to measurement errors and crude functional form, the census tract fixed effects in my main specification captured

[13] For model I, although the coefficient on the variable *Incre Dist_2nd City* Post 2008 Dummy* is not statistically significant, the Wald statistic for incremental distance to nearest city center in 2009 and 2010 is 0.52, with a p-value of 0.4714, which means the effect of second city center is no longer significant after 2008.

part of the effect of urban proximity, leading to a lower estimate of the urban premium. Thirdly, model VI reveals that there is no significant decline in urban influence in the year 2008 compared to 2001-2006, validating my assumption that there is a time lag before the housing market bust starting from early 2007 transmitted into related surrounding farmland markets. Finally, results of model VII reveal that there is no significant change in the effects of the most important influence variable *the distance to nearest urban center* if I assume the housing market bust happened in 2005. This supports the notion that there were no fundamental demand concerns other than the housing market bust in 2007 that could result in a downward trend in urban influences on farmland values since 2001.

	Whole sample		<10 miles		10-20 miles		30-60 miles	
	Boom	**Bust**	**Boom**	**Bust**	**Boom**	**Bust**	**Boom**	**Bust**
Total Urban Premium	**$1947**	**$1021**	**$2993**	**$1670**	**$2258**	**$1350**	**$1158**	**$669**
	($1086)	($579)	($1493)	($739)	($1006)	($635)	($465)	($281)
1) miles to nearest city center	$1374	$571	$2185	$951	$1631	$741	$721	$351
	($727)	($279)	($865)	($312)	($600)	($252)	($322)	($140)
2) incremental distance to second nearest city center	$284	$85	$255	$75	$268	$70	$308	$104
	($199)	($54)	($294)	($61)	($217)	($61)	($122)	($45)
3) surrounding urban population	$231	$368	$390	$662	$294	$541	$112	$215
	($231)	($320)	($328)	($404)	($246)	($399)	($95)	($140)
4) gravity index	$59	-$2	$165	-$17	$66	-$2	$17	-$1
	($93)	($39)	($183)	($133)	($66)	($2)	($12)	($1)
Number of observations	9079	1517	1293	128	2854	406	2044	478

Continued

Table 4. Comparison of Urban Premiums Before and After the Housing Market Bust – Model 0

Table 4 continued
Note: The values of *miles to nearest city center*, *incremental distance to second nearest city and gravity index* after 2008 are also included in the total value of the urban premium although their corresponding coefficients are not significant at 10% level. <10 miles, 10-20 miles, and 30-60 miles are the distance from a farmland parcel to nearest city center. Standard deviations are in parentheses.

To better understand the magnitude of the structural change, I use the regression results in Tables 2 and 3 to develop estimates of urban premiums in Tables 4 and 5 following the methods illustrated in section III.c. The four main urban influence variables are included in the construction of the urban premium even if their coefficients are statistically insignificant. From Table 4, I observe that, before 2007 relative to the reference parcel not subject to urban influence, the agricultural parcels subject to urban influence on average enjoy a $1,947 per acre urban premium, or roughly 43% of the per-acre sales prices (without structures). However, after 2008, a sizeable reduction in the urban premium occurred: it declined to only $1,021 per acre on average, which is about 23% of the average per-acre sales price.

I also find that, as expected, the urban premium is on average higher for parcels in closer proximity to urban centers (Table 4), and the impact of the residential housing market bust varied with urban proximity: the difference in the size of the urban premium for parcels within 10 miles of the nearest city center was around $1,835 greater than that for parcels at least 30 miles away from urban centers before 2007, on average, and this difference shrank to about $1,001 after the housing market bust[14]. In other words, the housing market bust has a greater impact on parcels closer to urban centers than those farther away, and resulted in some convergence of the size of the urban premium between these two groups. Also, previous studies have typically only considered the distance to nearest city center when measuring urban influence (Guiling, et al. 2009), yet comparison of Table 4 and Table 5 model II reveals that not accounting for the joint effects of proximity to multiple urban centers may significantly underestimate the size of the urban premium by as much as 17%, at least in periods of strong housing market growth: before 2007, the total urban premium would drop to $1,627 on average without three measures for multiple urban centers, including the *incremental distance to second nearest city center*, *surrounding urban population*, and the *gravity index*. This highlights the significant undervaluation of the effects of the urban influences when only *the distance to nearest city center* is included, which is common in previous studies.

Measures of urban premiums across different specifications shown in Table 5 are fairly robust: agricultural land parcels in all specifications experienced, on average, a significant decline in urban premium after the housing market bust, by more than half for models with census tract fixed effects. Although the absolute dollar value for the urban

[14] Alternative specifications of urban influences yield similar results: e.g. the urban premiums for parcels in MSA counties are about 1.5 times that for parcels in non-metropolitan counties, on average.

	Model I		**Model II**		**Model III**		**Model IV**		**Model V**		**Model VI**		**Model VII**	
	Boom	**Bust**	**Boom**	**Bust**	**Boom**	**Bust**	**Boom**	**Bust**	**Boom**	**Bust**	**Boom**	**2008**	**01-04**	**06-10**
Total Urban Premium	**$1993**	**$1136**	**$1627**	**$959**	**$1829**	**$826**	**$3379**	**$1685**	**$2273**	**$1675**	**$2056**	**$1899**	**$2016**	**$1745**
	($1127)	($693)	($810)	($420)	($1028)	($456)	($2292)	($1513)	($1111)	($670)	($1128)	($870)	($1127)	($728)
1) miles to nearest city center	$1417	$633	$1627	$959	$1296	$465	$2355	$978	$1730	$1079	$1509	$1734	$1430	$1403
	($770)	($367)	($810)	($420)	($694)	($219)	($1601)	($774)	($882)	($871)	($804)	($871)	($765)	($626)
2) incremental distance to second nearest city center	$282	$119			$262	$73	$511	$5.3	$487	$447	$290	$311	$309	$270
	($197)	($75)			($184)	($46)	($454)	($4.8)	($332)	($278)	($201)	($205)	($221)	($174)
3) surrounding urban population	$238	$387			$218	$290	$437	$710	$6.3	$151	$206	-$147	$227	$71
	($234)	($327)			($217)	($253)	($429)	($802)	($6.2)	($128)	($203)	($129)	($227)	($60)
4) gravity index	$56	-$2			$54	-$2	$76	-$8	$50	-$2	$51	$1.25	$49	$0.24
	($87)	($36)			($85)	($32)	($105)	($164)	($81)	($31)	($81)	($37)	($81)	($9)
Number of observations	9078	1517	9086	1477	9079	1517	8558	1513	9083	1517	9079	1262	6271	5445

Table 5. Robustness Checks of Predicted Urban Premium Across Different Hedonic Models

Note: The values of *miles to nearest city center*, *incremental distance to second nearest city*, *surrounding urban population and gravity index* are also included in the construction of urban premium although their corresponding coefficients are not significant at 10% level. Standard deviations are in parentheses.

premium is much higher for model IV, the total urban premium accounts for 45.8% of the prices with structures on average, which is consistent with model 0 using prices without structures. Consistent with previous discussions on the magnitude of the coefficients, model V with county fixed effects yields a much higher estimate of the urban premium. Model VI shows that in the year 2008, there is no evidence of significant decline in the urban influence and the proximity to nearest city center remains the most important contributor of the urban influence variables. In addition, model VII reveals that the urban premium stayed fairly constant before 2007[15], and the significant downward pressure was imposed by the housing market bust rather than other demand issues.

These results also reveal that there is rich spatial heterogeneity in the parcel-level measure of urban premium from one parcel to another: prior to 2007 the urban premium, with an average of $1,947 per acre (Table 4 whole sample), ranges from $145 per acre for parcels that are more than 50 miles away from the nearest city center to almost $8,000 per acre for parcels within urbanized areas. A map of estimated urban premiums based on the results of model 0 (Table 4) is included in Figure 4 in the following. This rich spatial heterogeneity of the urban premium suggests that even in Ohio where almost all parcels are subject to some degree of urban influence, the actual magnitude of the value of the urban influence varies substantially across space.

I previously described the potential for omitted variable bias arising from spatial dependence, as the land parcels in my data are spatially ordered. I tested for spatial autocorrelation using *Moran's I* test, where a positively significant *I* would indicate that

[15] Another robustness check using 2001 to 2004 as the pre period and 2006 to 2008 as the post period reveal that the average urban premium between 2006 and 2008 is $1584.

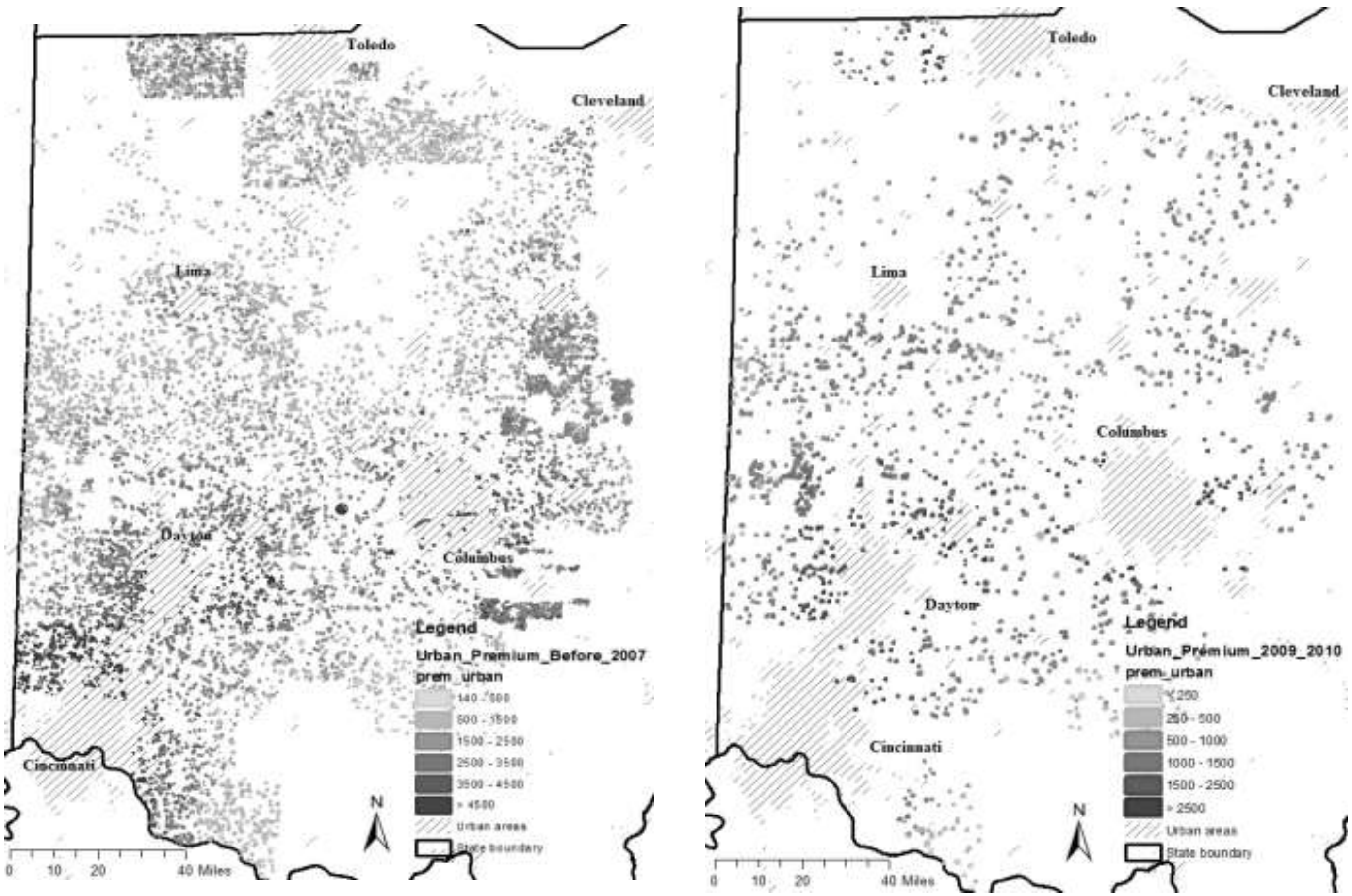


Figure 4. Spatial Distribution of the Urban Premium Before 2007 and After 2008

Model	(a)	(b)	(c)	(d)
Dist_City*within 10 miles	-0.0096***	-0.0092***	-0.0094***	-0.1300***
	(0.0015)	(0.0012)	(0.0013)	(0.0229)
Dist_City*within 10 miles*Post 2008 dummy	0.0038	0.0051**	0.0050***	0.0991**
	(0.0027)	(0.0026)	(0.0016)	(0.0492)
Dist_City*beyond 10 miles	-0.0102***	-0.0081***	-0.0087***	-0.1370***
	(0.0013)	(0.0011)	(0.0011)	(0.0218)
Dist_City*beyond 10 miles*Post 2008 dummy	0.0049*	0.0051**	0.0070***	0.1111**
	(0.0026)	(0.0026)	(0.0011)	(0.0472)
Incre Dist_2nd City	-0.0037***	-0.0038***	-0.0053***	-0.0252***
	(0.0008)	(0.0007)	(0.0007)	(0.0068)
Incre Dist_2nd City*Post 2008 dummy	0.0016	0.0038**	0.0082***	0.0123
	(0.0017)	(0.0017)	(0.0012)	(0.0159)
Urban population within 25 miles	0.0002***	0.0003***	0.0002***	0.0003***
	(5.13E-05)	(4.4E-05)	(4.51E-05)	(4.44E-05)
Urban popul within 25 miles	7.99E-05	0.0001	0.0002**	9.82E-05
*Post 2008 dummy	(0.0001)	(0.0001)	(8.41E-05)	(0.0001)
Gravity index	1.85E-05***	2.62E-05***	2.2E-05***	1.15E-05*
	(5.68E-06)	(5.65E-06)	(5.63E-06)	(6.46E-06)
Gravity index*Post 2008 dummy	-1.90E-05***	-2.70E-05***	-2.3E-05***	-1.20E-05*
	(5.86E-06)	(5.68E-06)	(5.66E-06)	(6.47E-06)
Building area % of parcel	0.0793	0.0961*	0.1112**	0.0592
	(0.0534)	(0.0518)	(0.0511)	(0.0523)
Distance to highway ramp	-0.0021	-0.0045	-0.0019	-0.0129***
	(0.0034)	(0.0033)	(0.0032)	(0.0050)
Distance to railway station	-0.0008	-0.0045	-0.0045	0.0006
	(0.0038)	(0.0036)	(0.0036)	(0.0086)
Year fixed effects	yes	yes		yes
Price deflator using quarterly Housing Price Index			yes	
Functional form	Log-linear	Log-linear	Log-linear	Log-log
Spatial fixed effects	Block group	Township	Census tract	Census tract
Root mean squared error	0.6170	0.6301	0.6200	0.6244
Adjusted R-square	0.2505	0.2216	0.2432	0.2324
Number of observations	10604	10604	10817	10604

Continued

Table 6. Additional Robustness Checks of Hedonic Regressions

Table 6 continued

Note: the dependent variable in this model is the log of per-acre agricultural land prices without structures. *, **, and *** indicates the coefficient is significant at 10%, 5% and 1% level, respectively. 505 census tract fixed effects are included in the model. Column (a) uses 1303 block group fixed effects instead of 505 census tract fixed effects, while column (b) uses 315 township fixed effects. Column (c) uses quarterly Housing Price Index from Federal Housing Finance Agency, while the other specifications just use year fixed effects without a price deflator. In column (d), a log-log specification is adopted where all proximity variables on the right hand side enter the regression in a logarithm form.

the variable value at each parcel tends to be similar to nearby neighbor parcels (Anselin and Hudak 1992). The global and local spatial autocorrelation by *Moran's I* test and the *Geary's C* test both indicated that although some explanatory variables are spatially correlated, the residuals from the hedonic regressions exhibit no patterns of spatial autocorrelation. The various measures of urban influences and agricultural productivity appear to adequately control for any inherent spatial correlation. Additional robustness checks using block group fixed effects shown in Tables 6 and 7 column (a) yield similar results as model 0, indicating that census tract fixed effects in my main specification could adequately control for omitted variables at the subcounty level.

The standard hedonic price method assumes linear parameterization and fixed functional form, which may introduce bias when the functional form for certain explanatory variables is not correct. To address this potential misspecfication bias, I ran two additional robustness checks. The first one adopts a log-log specification rather than the log-linear form used in all previous regressions, and the results are shown in Tables 6 and 7 column (d). The second one involves propensity score matching (PSM), which does not assume a particular functional form for the price function (Heckman and Navarro-Lozano 2004).

To implement matching, I constructed treatment and control groups based on distances to nearest city center, and ran several difference-in-difference regressions and regular regressions on the matched sample using different matching algorithms and different definitions of proximity to urban centers. Although the magnitude of urban premium is not the same, these two robustness checks both yield qualitatively similar conclusion as the main specification that the value of being close to urban areas significantly declined due to the recent housing market bust.

Conclusion

Because farm real estate values are such significant components of the farm sector balance sheets and farm household investment portfolios, understanding the key determinants of changes in U.S. farmland prices are of perennial interest to policymakers. Yet, little is known about how significant changes in competing land markets affect farmland values. With more than one-third of farmland estimated to be subject to urban influences, the effects of changes in demand for residential housing markets are of special interest. In particular, quantifying the effects of the housing market 'bust' offers unique insights into the dynamics of the relative importance of different determinants of farmland values, and helps inform on the linkages between urban and rural land markets. By controlling for spatial heterogeneity using localized fixed effects and developing a parcel level measure of "urban premium" (the value attributable to urban demands for

	(a)		(b)		(c)		(d)	
	Boom	**Bust**	**Boom**	**Bust**	**Boom**	**Bust**	**Boom**	**Bust**
Total Urban Premium	**$1927**	**$1363**	**$1985**	**$906**	**$1931**	**$680**	**$1261**	**$718**
	($1177)	($743)	($1089)	($637)	($1073)	($698)	($948)	($539)
1) miles to nearest city center	$1404	$874	$1355	$489	$1301	$492	$689	$139
	($849)	($471)	($721)	($330)	($720)	($360)	($516)	($100)
2) incremental distance to second nearest city center	$292	$216	$304	-$5	$376	-$324	$158	$96
	($217)	($148)	($206)	($3)	($264)	($200)	($156)	($88)
3) surrounding urban population	$182	$275	$256	$424	$203	$515	$374	$487
	($189)	($250)	($239)	($353)	($198)	($402)	($377)	($427)
4) gravity index	$50	-$1	$70	-$2	$52	-$3	$40	-$3
	($81)	($21)	($107)	($30)	($82)	($41)	($64)	($51)
Number of observations	9071	1517	8902	1476	9190	1621	9082	1517

Table 7. Predicted Urban Premium Across Additional Robustness Checks in Table 6
Note: standard deviations in parenthesis

developable land), this study provides the first concrete evidence of a decline in the effect of urban influences on surrounding farmland values shortly after the housing market bust. Using a hedonic modeling approach and farmland parcel sales data in Western Ohio, this study estimates the magnitude of the urban premium at $1,974 per acre on average before 2007 and $1,021 per acre on average in 2009 and 2010, a reduction in the value of proximity to urban areas from more than 40 percent to about 20 percent of farmland values as a result of the residential housing market bust. In other words, farmland was not immune to the residential housing bust; the portion of farmland value attributable to urban demands for developable land was almost cut in half shortly after the housing market bust in 2009-2010. My results also demonstrated that not accounting for multiple urban centers using variables such as proximity to the second nearest city center can underestimate the contribution of the urban premium by as much as 17 percent, at least in periods of strong residential housing market growth. In addition, compared to models with fine-scale, local spatial fixed effects, models with county fixed effects tend to lead to a less conservative estimate of the urban premium. By removing the effects of time-invariant omitted variables at the census tract level to achieve unbiasedness, my main specification relies on the variation within census tracts, which could lead to an underestimate of the total effect as pointed out by (Abbott and Klaiber 2011). In other words, the urban premium shown in my paper could be just a partial effect. Furthermore, my analysis shows that matching, although free from the functional form specification bias, doesn't add much to my analysis.

Despite the decline in the significance and magnitude of the urban premium after 2008, farmland prices remained relatively steady over my study period - a trend that is comparable to previous studies that suggest rural housing values have declined more than farm real estate (land with structures) in most U.S. states as a result of the urban residential housing market bust (Nickerson, et al. 2012).[16] Increased commodity demands over this period appear to have contributed to the trend; the significant effect of *proximity to ethanol plant after 2008*, for example, indicates that proximity to new commodity buyers may have substantially obscured the impact on farmland values of the downturn in the urban residential housing market. my findings of a significant decline in the impacts of urban influences in 2009 and 2010 are short-run effects, and do not necessarily suggest urban influences no longer matter for surrounding farmland parcel values in the long run. An analysis of the long-term impacts of the housing market bust on surrounding farmland values would require many additional years of data and thus is beyond the scope of this study.

[16] The number of sales of farmland parcels in Western Ohio dropped by 40% as a result of the housing market bust from an average of 1502 annually between 2001-2006 to 904 on average over 2008-2010.

Chapter 2: The Expanding Ethanol Market and Farmland Values: Identifying the Changing Influence of Proximity to Agricultural Market Channels

Introduction

Farmland values represent over 80 percent of the value of the farm sector assets, and farmland represents the largest asset in the typical farm household investment portfolio (Nickerson, et al. 2012). As a result, U.S. farmland values and the factors influencing these values have long been of the subject of a great deal of economic research (Nickerson, et al. 2012). With the strong federal support represented by the Energy Policy Act of 2005 and the Energy Independence and Security Act of 2007, the number of ethanol plants saw a four-fold increase along with a dramatic increase in U.S. ethanol production, making U.S. the largest ethanol producer in the world. These macroeconomic trends in the ethanol market are speculated to have elevated agricultural commodity prices, farmers' expectations about future profits, and farmland values (Low and Isserman 2009; Wallander, et al. 2011). The growing biofuels market, along with other factors such as historically low interest rates (Schnitkey and Sherrick 2011) and rising demand for U.S. grain exports (Gloy, et al. 2011), helped fuel the recent remarkable rise in farmland values in Corn Belt states. The steady increase in rural farmland values in the Corn Belt states indicates that these macroeconomic shifts in agricultural markets seem significant enough to offset the downward pressure resulting from the recent residential housing market bust (Zhang and Nickerson forthcoming). Understanding how farmland values respond to these agricultural market changes, and in particular to expanding ethanol production, is of critical policy interest given the importance of farmland values for the farm sector and farm household wellbeing.

Numerous studies have analyzed the farm and non-farm determinants of farmland values, including soil quality measures (Huang, et al. 2006; Palmquist and Danielson 1989), urban proximity (Livanis, et al. 2006; Shi, et al. 1997), environmental amenities (Bastian, et al. 2002), wildlife recreational opportunities (Henderson and Moore 2006), zoning (Chicoine 1981), and farmland protection easements (Nickerson and Lynch 2001). In contrast, evidence of the potential impact of access to agricultural market channels, such as proximity to ethanol plants or grain elevators, is limited. Despite some recent research on the impacts of the ethanol industry on crop prices (Gallagher 2006; McNew and Griffith 2005), most previous studies of farmland markets have not considered the influence of agricultural market variables such as proximity to ethanol plants. Of the several studies (Blomendahl, et al. 2011; Henderson and Gloy 2009; Nehring, et al. 2006) that have considered the capitalization effect of proximity to ethanol plants on farmland values, all employ the standard hedonic price model, which, despite its popularity, suffers from a number of limitations in terms of identification (Bajari, et al. 2012). In particular, the location of an ethanol plant is a non-random process affected by surrounding locational features such as the availability of feedstock nearby and the access to road networks (Lambert, et al. 2008). As a result, the estimates from a simple hedonic model may suffer from sample selection bias due to systematic differences between those parcels that are located nearer versus farther away from an ethanol plant (Imbens and Wooldridge 2009). The measure of proximity to nearest ethanol plant that is commonly used in previous studies may also be endogenous due to the non-random site selection process of ethanol plants.

The study closest in spirit to this study is Towe and Tra (2012), which used a difference-in-difference propensity score matching estimator to quantify the effect of the 2005 ethanol mandate into farmland values. They find that new ethanol facilities had no effect on nearby farmland values prior to the mandate (2002-2004) but had significant effects after the policy (2004-2006). However, there are a number of important differences that distinguishes my work. First, Towe and Tra (2012) aim to examine average effect of the 2005 federal ethanol mandate with a focus on its creation of exuberant confidence in the expected farmland returns beyond market fundamentals, while this study seeks to quantify the spatial-explicit capitalization of new ethanol plants in surrounding farmland values due to reduced . Second, Towe and Tra (2012) used farmer-reported survey data on land values while I use actual arms-length sales records of farmland parcels. Finally, the longer span from 2001 to 2010 in my dataset, as opposed to 2002-2006 in Towe and Tra (2012), allows me to empirically investigate the role of ethanol sector during the recessionary time due to the recent housing market bust.

The objective of this study is to identify the marginal value of proximity to ethanol plants and other agricultural market channels and to test for structural change in these effects before and after the ethanol market expansion in Ohio in late 2006-early 2007. I hypothesize that changes in agricultural output markets, including increased demand for biofuels and grain exports, were capitalized into agricultural land values and thus a greater influence of proximity to ethanol plants after their developments is expected. This study uses parcel-level data on agricultural land sales from 2001 to 2010, a period which encompasses the expansion of ethanol facilities in Ohio, for a 50-county region of Ohio that encompasses the great majority of grain production in Ohio.

I address the aforementioned sample selection bias and potential endogeneity of proximity to ethanol plants using both matching and instrumental variables (IV) approaches. Specifically, for each of the three types of agricultural market channels, I use propensity score matching (PSM) to construct a matched sample that controls for systematic differences in observable characteristics between parcels that are within close proximity to these destinations and those located farther away. For grain elevators and agricultural output terminals, standard hedonic regressions on the matched sample are used to test for the relative effect of proximity to these destinations has changed due to the constructions of ethanol plants. Matching alone, however, does not address the potential endogeneity of the location of ethanol plants. Instead, I use instrumental variables regression on the matched sample for ethanol plants to test for the effects of proximity to newly constructed ethanol plants To control for this endogenous proximity measure, I construct two instruments that are based on the idea of spatial competition among agricultural market channels. Specifically, given the significance of transportation costs in the value of agricultural commodities (Fackler and Goodwin 2001), a new ethanol plant should find it optimal to locate a certain distance away from other agricultural markets in order to minimize competitive pressure and maximize their market area. With this in mind, I construct two instruments: capacity weighted average distances to other, non-nearest ethanol plants and capacity-weighted distances to other agricultural output terminals. These instruments, which capture the competitive pressure faced by a particular ethanol plant, would affect the site selection of this plant and thus the distances from it to farmland parcels, but would not directly impact the value of farmland parcels closer to this plant since effects of proximity to ethanol plants are relatively local (Gallagher 2006).

The main result provides evidence for positive and significant marginal value of being within close proximity to an ethanol plant following construction of seven ethanol plants in or near western Ohio in late 2006-early 2007. Specifically, results from the instrumental variables estimation with the matched sample suggest that the marginal value of farmland increases by $46 per mile per acre within proximity to the nearest ethanol plant following construction of these plants. By comparison, the effect of proximity to nearest city center and second nearest city is $30-66 and $30-40 per mile per acre, respectively. Results also reveal a stronger influence of proximity to grain elevators as well as a reduction in the magnitude and significance of the effect of proximity to agricultural terminals after early 2007 due to competition from the newly constructed ethanol plants. Specifically, I find that the marginal value of being close to an agricultural terminal reduces from $48 to $30 per mile per acre after early 2007. These results demonstrate the growing importance of the biofuels market for farmland values. A comparison between the standard hedonic estimates and the instrumental variables estimates confirms the endogeneity of proximity to ethanol plants, which, if left uncontrolled for, would result in a downward bias in the standard hedonic estimates due to unobserved characteristics.

This study makes an important contribution to the literature on farmland valuation and the policy debate about the welfare effects of ethanol market expansion. While the ethanol market expansion results in elevated commodity prices, farmer income, and farmland values as shown in this study, it has been criticized for its high dependency on government subsidies and potential negative impacts on environmental quality through its incentives for corn expansion (Cappiello and Apuzzo 2013; Tiffany 2009). To the best of my knowledge, this study is the first to provide formal evidence of the effects of ethanol market expansion on farmland values during a strong recessionary time that exerted substantial downward pressure – the common wisdom that the rise of ethanol industry has helped the farm sector withstand the downturn (Nickerson, et al. 2012). Second, by combining matching and instrumental variables approaches, this study directly addresses the potential endogeneity of the proximity of farmland parcels to ethanol plants and is thus subject to less bias than the commonly used hedonic estimates, which typically yields a much lower capitalization effect (Blomendahl, et al. 2011; Henderson and Gloy 2009).

Theoretical Framework

Among the most influential theories that help explain the value of land is Ricardo's economic theory of rent (Ricardo 1996). Ricardo's key insight was that land which differs in quality and which is limited in supply generates rents that arise from the productive differences in land quality or in differences in location. Farmland value is comprised of the net present value of economic returns to land. The capitalization formula is written as

$$V_{it} = E_t \sum_s \frac{R_{is}}{(1+\delta_t)^{s-t}}, \quad where\ s = t, t+1, \ldots \qquad (9)$$

In this formulation, the value of agricultural land parcel i at time t V_{it} is defined as the expected future annual returns to farmland R discounted at rate t. Any factor affecting the farmland returns R, either in terms of agricultural productivity, recreational opportunities or potential profitability of development for urban uses, would impact the farmland values. Formally, the farmland returns R_{it} can be approximated by a linear combination of parcel attributes and location characteristics $\boldsymbol{X_{it}}$ using Taylor expansion; a common linear specification is defined as

$$R_{it} = \beta' \boldsymbol{X_{it}} + \tau_t + \eta_{it} \qquad (10)$$

where τ_t is time fixed effects and η_{it} is the remaining normally distributed error term.

The vector of parcel attributes and location characteristics $\boldsymbol{X_{it}}$ can be further decomposed into four categories: (1) the parcel-specific agronomic variables $\boldsymbol{A_{it}}$ such as soil quality and slope of the parcel; (2) the natural amenities variables $\boldsymbol{N_{it}}$ such as varied topography and proximity to surface water; (3) the urban influence variables $\boldsymbol{U_{it}}$ such as surrounding urban population and access to highway; and (4) the newly emerging set of agricultural market influence variables $\boldsymbol{M_{it}}$ such as proximity to ethanol plants, grain elevators, and agricultural product terminal ports, so that

$$\boldsymbol{X_{it}} = \boldsymbol{A_{it}} + \boldsymbol{N_{it}} + \boldsymbol{U_{it}} + \boldsymbol{M_{it}} \qquad (11)$$

Therefore I get the following model specification:

$$V_{it} = E_t \sum_s f(\boldsymbol{A_{is}}, \boldsymbol{N_{is}}, \boldsymbol{U_{is}}, \boldsymbol{M_{is}};\ \delta_t)\ , where\ s = t, t+1, \dots \qquad (12)$$

The study region - western Ohio - is fairly homogenous in soil type, slope of the land, climatic conditions and surrounding land uses. As a result little variation in generating recreational income is expected among all the parcels. Hence the urban influence variables $\boldsymbol{U_{it}}$ and the agricultural market influence variables $\boldsymbol{M_{it}}$ are of particular interest. Agricultural land closer to urban fringe could sell for a premium, an option value that equals to the expected returns from the conversion into urban development at a future date (Capozza and Helsley 1989). The recent Great Recession may greatly diminish the urban option conversion value of the agricultural land, and as a result, a declining relative significance of the urban influence variables $\boldsymbol{U_{it}}$ in determining the farmland values is expected.

At the same time, much has changed in terms of agricultural market influence variables. Most notably, ethanol has been embraced enthusiastically as a promising alternative renewable energy (Low and Isserman 2009). Federal energy policies supporting the production of biofuels have increased demand for corn, which elevated corn and other agricultural commodity prices (Nickerson, et al. 2012). Previous studies have identified increased corn basis prices in the vicinity of an ethanol plant (McNew and Griffith 2005), which could translate into higher farmland values through capitalization. This increased demand, in part met by the supplies from local grain elevators, could also enhance the positive impact of the proximity to the grain elevators on farmland values. By attracting corn supplies from surrounding land parcels or nearby grain elevators, the new ethanol plants may constitute a competing source of demand for grains for traditional terminal markets (Nickerson, et al. 2012). However, whether the competition from ethanol plants is strong enough to offset the benefits of increased grain exports to China for the agricultural terminal markets is an empirical question.

Econometric Challenges and Empirical Strategy

The Identification Problem in the Hedonic Price Estimation

With Rosen's (1974) seminal work as a backdrop, the hedonic price method has become the workhorse model for valuing local public goods and environmental amenities (Bishop and Timmins 2011). Specifically, hedonic regression is the most commonly used approach for estimating the impact of environmental amenities and disamenities on real estate or land values (see Hite, et al. (2001); Kohlhase (1991); Palmquist (1989) for applications and Palmquist (2005) for a comprehensive review). Almost all of aforementioned literature on farmland valuation has employed the land value hedonics model. A common specification is the linear form defined as

$$V_{it} = \beta_0 + \beta_A{}'\boldsymbol{A_{it}} + \beta_U{}'\boldsymbol{U_{it}} + \beta_R{}'\boldsymbol{R_{it}} + \beta_M{}'\boldsymbol{M_{it}} + \tau_t + \varepsilon_{it}, \quad (13)$$

where the agricultural land values V_{it} are approximated by the nominal sale prices of the agri-cultural land without structures P_{it}. In this setting, agricultural land is regarded as a differentiated product with a bundle of agricultural-quality and location characteristics, and each characteristic is valued by its implicit price (Nehring, et al. 2006; Rosen 1974). Despite its popularity, the hedonic pricing method suffers from a number of well-known econometric problems (Bajari, et al. 2012; De Vor and De Groot 2011). In my settings, on one hand, the location selection of an ethanol plant is a non-random process affected by the availability of feedstock nearby, the access to navigable rivers, highways, or railroads, the access to sewer service and natural gas pipeline, and the extent of the product markets (Lambert, et al. 2008), which could lead to a upward bias for the hedonic estimates. Arguably, agricultural parcels closer to the ethanol plants could have better soil quality and easier access to the transportation network than that of those parcels further away. On the other hand, areas with low corn basis levels or low competitive pressure may also be chosen as sites for ethanol plants to minimize the land purchase costs (Towe and Tra 2012) or to minimize spatial competition with other agricultural markets, which would lead to a downward bias in the hedonic estimates. As a result, the estimates from a simple hedonic model have two interrelated econometric problems: first, with no control for systematic differences between those parcels that are located nearer versus farther away from an ethanol plant may be biased due to the unequal distribution of the covariates across the treatment and control subsamples (Imbens and Wooldridge 2009). Second, due to the non-random site-selection process of ethanol plants, the distance from farmland parcels to nearest ethanol plant may also be endogeneous.

Quasi-Experimental Design

I address the potential sample selection bias and endogeneity of distance to ethanol plant by employing a two-tiered quasi-experimental design which combines matching and instrumental variables (IV) approaches. Specifically, I use matching to address the sample selection on observables, and then use two instruments based on the idea of spatial competition to address the residual endogeneity of the proximity to ethanol plants, recognizing the potential limitations of matching in dealing with bias resulting from unobservables.

Matching is an increasingly popular procedure to control for sample selection bias due to observables (Imbens and Wooldridge 2009; Zhao 2004), which selects treated observations and control observations with similar characteristics, by covariates (Rubin 1980), or by propensity score (Rosenbaum and Rubin 1983). Zhao (2004) finds that propensity score matching (PSM) has smallest bias and performs well for a relatively large sample given his specific design. With a sample size of over 10,000, this study, as a result, uses PSM as the main matching technique. For each of the three types of agricultural market channels, I use PSM to construct treatment and control subsamples based on their proximity to these agricultural markets and then use regressions on the matched sample to test for the hypothesis that the relative effect of proximity to these destinations has changed since the construction of all seven ethanol plants in late 2006-early 2007. For grain elevators and agricultural output terminals, standard hedonic regressions on the matched sample are used, which, however, would yield biased estimates for ethanol plants due to their non-random site-selection problem. To address the residual endogeneity of proximity to ethanol plants, I construct two instruments that are based on the idea of spatial competition among agricultural market channels and estimate instrumental variables regressions on the matched sample instead.

Propensity Score Matching

The propensity score matching (PSM) estimator, as a way to identify average treatment effect, has gained popularity in land use and agricultural economics literature in recent years (Bento, et al. 2007; Lynch, et al. 2007). This study's application involves a two-step matching strategy. First, it trims the pre-construction sample to remove extremely dissimilar parcels with those sold after the month of construction of the nearest ethanol plant using one to four (oversampling) nearest neighbor matching technique suggested by Caliendo and Kopeinig (2008). Specifically, the propensity score (the probability of being sold after the month of construction of the nearest ethanol plant given observed parcel attributes $\mathbf{X_{it}}$) is calculated using a logit model in which treatment is modeled as a function of parcel attributes and other location characteristics, including parcel size, soil suitability, proximity to nearest employment center, proximity to nearest highway, surrounding land uses, and neighborhood population density. I further dropped parcels with propensity scores not on common support and those with propensity scores greater than 0.9 or lower than 0.05, as recommended by Crump, et al. (2009). Similar sample trimming technique has been used by Busso, et al. (2013).

With the trimmed sample, I then construct the treatment and control subsamples based on the proximity to a given type of market channels. Specifically, with *a priori* defined cutoff value for distance which defines proximity, parcels that are located within this cutoff distance to agricultural markets are assumed to be in the treatment group. These parcels are then matched with those parcels located farther away than the cutoff distance following the one to four nearest neighbor PSM technique illustrated above. The cutoff distances used to define proximity to ethanol plants, grain elevators, and agricultural output terminals are 10 miles, 5 miles, and 15 miles, respectively. These cutoff values are determined using semiparametric regressions and the covariate imbalance tests.[1]

The validity of the PSM approach hinges on the assumption that the model specification is correct and all relevant conditioning variables have been included in the PSM model (Diamond and Sekhon 2013). As a result, to ensure the main results do not depend crucially the particular matching methodology, I conduct multiple robustness checks by changing the matching algorithms (e.g., using covariate matching using Mahalanobis metric (Rubin 1980)), altering the parameters of PSM (use one or two nearest neighbors instead of four), or employing longer or shorter cutoff distances used to define proximity to agricultural market channels.

Instrumental Variables Regressions on the Matched Sample

Following Imbens and Wooldridge (2009) and for each of the three matched samples, regressions on the matched samples are used to test for a structural change in the influence of proximity to a given type of market channel before and after construction of the ethanol plant. Through matching, this study constructed the counterfactual control subsample for each type of agricultural market channel, which differs from the treatment subsamples only in the proximity to this type of market channel. For grain elevators and agricultural output terminals, standard hedonic regressions on the matched samples are used to test for the relative effect of proximity to these destinations has changed due to constructions of ethanol plants. In particular, I use the following specification:

$$P_{it} = \beta_0 + \beta_A{}'\boldsymbol{A_{it}} + \beta_U{}'\boldsymbol{U_{it}} + \beta_R{}'\boldsymbol{R_{it}} + \beta_M{}'\boldsymbol{M_{it}} + \beta_{M\_POST}{}'\boldsymbol{M_{it}} * D\_POST + \tau_t + \varepsilon_{it}, \tag{14}$$

where D_POST is a binary time dummy indicating that the parcel is sold after the month of construction of nearest ethanol plants. The coefficient, β_M, on variables like distances to nearest grain elevators or agricultural terminals captures the capitalization effects of proximity to these destinations before late 2006-2007, while β_{M_POST}, the coefficient on the interaction term between these proximity variables and the time dummy, represents the significance and magnitude of the structural change in their effect.

For the matched sample constructed based on proximity to ethanol plants, standard hedonic estimates could be biased due to the potential endogeneity of the location of ethanol plants. The endogeneity mainly results from the fact that the location of ethanol plants are more likely to be in areas with abundant corn supply or better soil quality, which would also affect the value of these neighboring farmland parcels. However, personal communications with managers of all seven ethanol plants in or near western Ohio reveal that abundant corn supply is only one factor in the site-selection process; other equally important factors include the access to highway and railway, and access to sewer service and natural gas pipeline. An ideal location of an ethanol plant would require that all these factors are satisfied, which rules out remotely rural area without sewer system and leaves me towns and villages as candidate sites. Figure 10 in the Appendix A plots the percentage of corn acreage within 50 miles for all towns with access to both railways and natural gas pipelines. It reveals that after controlling for these factors and beyond a threshold on county-level corn production, there is no systematic correlation between corn supply and the location of ethanol plants, which suggests that the endogeneity problem of the proximity to ethanol plants may not be serious. Nonetheless, this anecdote evidence does not clear the endogeneity concern. To control for the potential residual endogeneity of the proximity to the nearest ethanol plant, I construct two instruments based on the idea of spatial competition among agricultural market channels. Previous studies have shown that transportation costs account for a significant fraction of the value of the agricultural commodities (Fackler and Goodwin 2001). Due to existence of transportation costs, a standard result from the models of spatial competition is the principle of maximum differentiation: each firm has an incentive to locate farther away from its rivals to avoid price competition (d'Aspremont, et al. 1979). Specifically, transportation costs imply a new ethanol plant should find it optimal to locate a certain distance away from other agricultural market channels in order to maximize their market area. With this in mind, I construct two instruments: capacity weighted average distances to other, non-nearest ethanol plants and capacity-weighted distances to other agricultural output terminals. A negative correlation between these two instruments and the endogenous distance to nearest ethanol plant variable is expected due to the spatial competition. Similar instruments are used in the urban economics literature: for example, in a location sorting model, Bayer and Timmins (2007) used the fixed attributes of other locations as instruments for the share of individuals who choose a particular location as the exogenous attributes of other locations influence the demand for the specific location via the sorting equilibrium. These instruments, which capture the competitive pressure faced by a particular ethanol plant, would affect the site selection of this plant and thus the distances from it to farmland parcels, but would not directly impact the value of farmland parcels closer to this plant since effects of proximity to ethanol plants are relatively local (Gallagher 2006). In other words, an instrumental variables regression on the matched sample rather than a standard hedonic regression is estimated to test for the effects of proximity to newly constructed ethanol plants. Specifically, I employ a two stage least squares approach and estimate the following equations:

$$\boldsymbol{M_{it}} = \beta_0 + \beta_A{}'\boldsymbol{A_{it}} + \beta_U{}'\boldsymbol{U_{it}} + \beta_R{}'\boldsymbol{R_{it}} + \pi_Z{}'\boldsymbol{Z_{it}} + \tau_t + e_{it}, \quad \text{(15a)}$$

$$\boldsymbol{M_{it}} * D\_POST = \beta_0 + \beta_A{}'\boldsymbol{A_{it}} + \beta_U{}'\boldsymbol{U_{it}} + \beta_R{}'\boldsymbol{R_{it}} + \pi_Z{}'\boldsymbol{Z_{it}} + \tau_t + \epsilon_{it}, \quad \text{(15b)}$$

$$P_{it} = \beta_0 + \beta_A{}'\boldsymbol{A_{it}} + \beta_U{}'\boldsymbol{U_{it}} + \beta_R{}'\boldsymbol{R_{it}} + \beta_M{}'\widehat{\boldsymbol{M_{it}}} + \beta_{M\_POST}{}'\widehat{\boldsymbol{M_{it}} * D\_POST} + \tau_t + \varepsilon_{it},$$

(15c)

where $\boldsymbol{Z_{it}}$ are these two instruments, and $\widehat{\boldsymbol{M_{it}}}$, $\widehat{\boldsymbol{M_{it}} * D\_POST}$ are the predicted values from the first-stage regressions.

Unlike the grain elevators and agricultural terminals which existed throughout the 2000s decade, the ethanol plants in or near western Ohio all started construction in late 2006 – early 2007. As a result, the coefficient on the distance to ethanol plants variable β_M has no intuitive interpretation, while β_{M_POST} captures the significance and magnitude of the spatial effects of proximity to ethanol plants following construction of these plants. To address the potential capitalization effects of proximity to an ethanol facility before its construction due to expectations, I run multiple robustness checks which include parcels sold six months or one year before the construction of their nearest ethanol plant in the post period. By controlling for sample selection bias and potential endogeneity of ethanol plant locations, my estimator is subject to less bias than the standard hedonic estimates (Imbens and Wooldridge 2009).

Data

Western Ohio hosts a vast majority of the state's agricultural land and provides an excellent laboratory to study the structural change in the proximity to agricultural market channels on farmland values in the context of ethanol market expansion. The biofuels industry in Ohio is gaining momentum over the last decade, with seven ethanol plants started construction in or near western Ohio in late 2006 – early 2007. I assembled a detailed database of 21,342 arm’s length agricultural land sale records for 50 counties in or near western Ohio from 2001 to 2010 obtained from the U.S. Department of Agricultural Economic Research Service and merged with purchased sales data from a private firm, CoreLogic. I now briefly describe the key elements of the data in additional detail.

To form the dataset of agricultural land transactions, this study combines the dataset (29 counties) purchased from CoreLogic, with the data from USDA ERS data (14 counties) and the data collected from county auditor office for counties like Seneca, Hardin, Allen, Lucas, Auglaize, Henry and Hamilton in Ohio and Randolph County, Indiana.[2] Only those agricultural parcels sold between 2001 and 2010 and with a valid arms-length indicator[3] are kept. Those valid agricultural sale records are merged with GIS parcel boundaries or are geocoded based on property addresses using Google Maps API. The sales prices are adjusted for the value of the structures on the farmland. Specifically, the

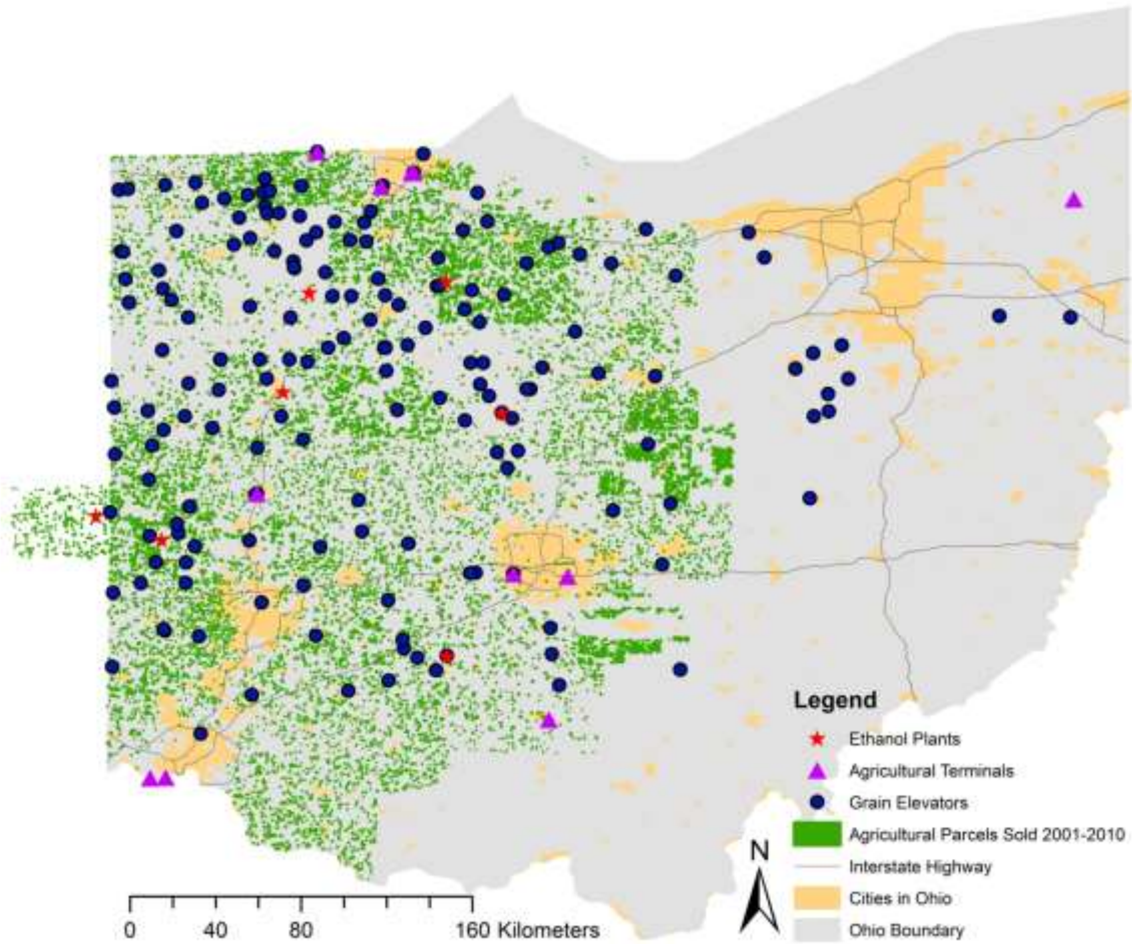


Figure 5. Agricultural Land Sales 2001-2010 and Agricultural Market Channels in Western Ohio

new sales prices are calculated as a fraction of the original prices, with the ratio being the percentage of assessed values of land only over assessed values of land and buildings altogether. Parcels with sales prices above $20,000/acre or below $1,000/acre are

	Unit	Mean	Std. Dev.	Min.	Max.
General Parcel Attributes					
Sale price	Dollars	143175	244364	303.7112	1.17E+07
Sale price per acre	Dollars	4362.70	3644.27	1000.161	19988.09
Assessed land value	Dollars	74190.62	162980	0	5878840
Assessed improvement value	Dollars	32269.28	68910	0	3937580
Assessed land value % of total assessed	%	0.7224	0.312	0.0035	1
Total acres	Acres	44.204	61.289	0.14	2380.66
Sale year	Year	2004.838	2.761	2001	2010
Agricultural Productivity Variables					
NCCPI	Number	5778.153	1518	0	8800.8
Cropland % of parcel	%	0.5502	0.3715	0	1
Soil class 1 area % of parcel	%	0.2810	0.3235	0	1
Soil class 2 area % of parcel	%	0.0778	0.1824	0	1
Soil class 3 area % of parcel	%	0.4406	0.4175	0	1
Steep slope (>15 degrees)	Binary	0.1888	0.3913	0	1
Urban Influence Variables					
Building area % of parcel	%	0.0364	0.1314	0	1
Distance to urban area of over 25k people	km	19.12	12.76	0	56.82
Total urban population within 25 miles	Thousand	290.04	231.94	64.7721	1187.38
Distance to highway ramp	Km	5.2218	3.29	0	19.10
Distance to nearest city	Km	40.9684	20.34	0.1983	105.66
Distance to nearest railway access point	Miles	3.1390	1.8080	0.005	11.254
Gravity index using three nearest cities	Number	1072.77	34101	52.76	4255332

Continued

Table 8. Summary Statistics of Agricultural Land Sales 2001-2010 in Western Ohio

Table 8 continued

Agricultural Market Influence Variables					
Distance to nearest ethanol plant	Km	45.99	22.57	0.68	111.75
Production capacity of nearest ethanol plant	Mgal	88.56	25.02	54.00	120.00
Number of ethanol plants within 25 miles	Number	1.13	0.91	0.00	4.00
Total production capacity of ethanol plants within 25 miles	Mgal	96.14	76.48	0.00	304.00
Distance to nearest grain elevator	Km	13.45	11.35	0.04	88.43
Distance to nearest agricultural terminal	Km	52.50	22.74	0.20	119.40
Capacity-weighted distance to other ethanol plants	Km	71.64	15.92	38.76	111.23
Capacity-weighted distance to other AG terminals	Km	101.49	45.70	7.21	204.54
Environmental Amenities Influence Variables					
Forest area % of parcel	%	0.153	0.259	0	1
Wetland area % of parcel	%	0.003	0.029	0	1
Pasture area % of parcel	%	0.120	0.241	0	1
Open water % of parcel	%	0.003	0.024	0	0.746
Observations			16434		

dropped along with parcels sold in the year 2007. Figure 5 shows a plot of the filtered sample consisting of 13,865 valid transactions. As is evident from the figure, these data are widely distributed over virtually the entire region. The locations of three sets of agricultural market channels - ethanol plants, grain elevators and agricultural terminal ports - are also shown in Figure 5.

Data on parcel attributes and location characteristics were obtained largely from the U.S. Department of Agriculture Natural Resources Conservation Services GeoSpatial Data Gateway, including the Census TIGER/Line Streets, National Elevation Dataset, National Land Cover Dataset, Soil Survey Spatial Data (SSURGO). Additional data on locations of cities and towns in Ohio was obtained from Ohio Department of Transportation (2012). I also used Census Block Shapefiles with 2010 Census Population and Housing Unit Counts (U.S. Census TIGER/Line 2012) to calculate the surrounding urban population. Data on ethanol plants, grain elevators and agricultural terminal ports were obtained from the Ohio Ethanol Council (2012), Farm Net Services (2012) and Ohio Department of Agriculture (2012). Using these data and ArcGIS software, I was able to create the parcel attributes and location characteristics vector $\boldsymbol{X_{it}}$. See Table 8 for summary statistics.

Most of variables in Table 8 are self-explanatory; however, three remarks need to be made. First, the variable National Commodity Crops Productivity Index is an interpretation in the National Soil Information System (NASIS). Specifically, the interpretation uses natural relationships of soil, landscape, and climate factors to model the response of commodity crops (see Dobos, et al. (2008) for details). Secondly, soil class 1 is defined as "All areas prime farmland", class 2 as "Prime farmland if drained", class 3 as "Farmland of local importance" and class 4 as "not prime farmland". Finally, this study highlights the set of the urban influence variables $\boldsymbol{U_{it}}$ and the agricultural market influence variables $\boldsymbol{M_{it}}$ in particular. Three aspects of urban influences are considered: distance to nearest city captures the importance of urbanized areas as commuting hub or sources of non-farm income, proximity to urbanized areas and road network and surrounding urban population represent the option value of future land conversion to urban uses. Surrounding urban population also captures the consumer demand for agricultural products, which will drive up the agricultural returns. Proximity variable for each of the three agricultural market channels is calculated as driving distance from farmland parcels to nearest market.

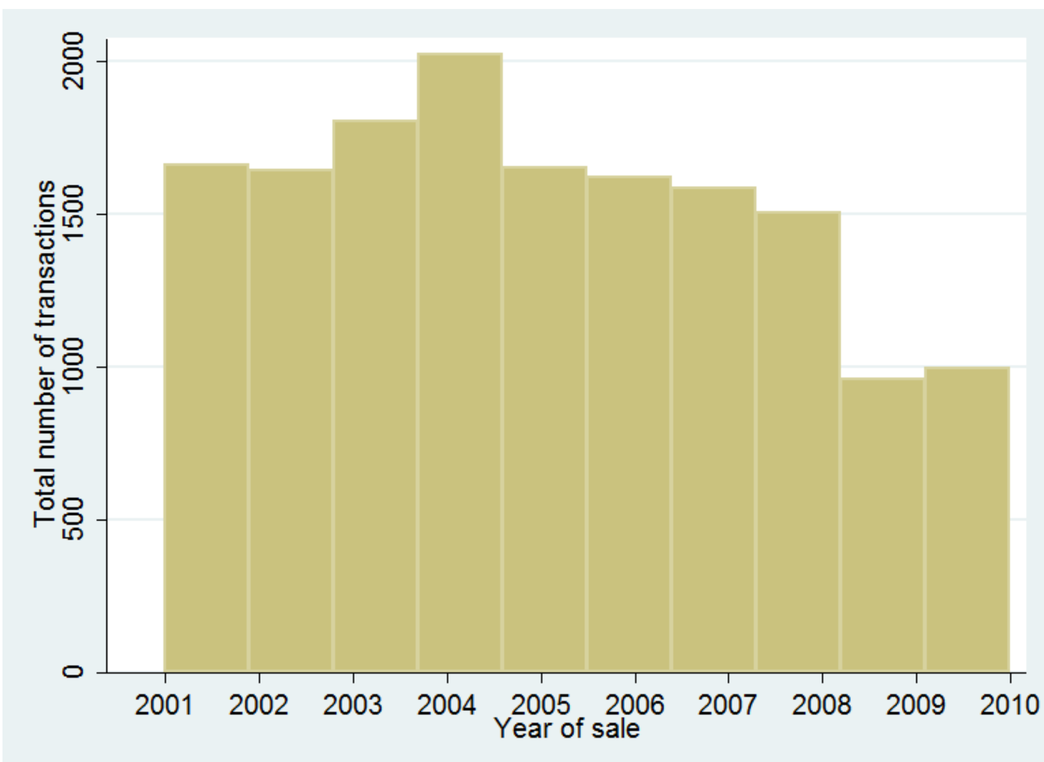


Figure 6. Number of Agricultural Land Sales 2001-2010 in Western Ohio

Results and Discussion

The recent Great Recession have led to a dramatic decline in urban land and housing values across the U.S. The same is not true, however, of agricultural land values. Figure 6 and 2 plot the number of agricultural land sales and the average farmland values in western Ohio since 2001, respectively. Although the number of farmland sales dropped precipitously after the housing market bust, there was no corresponding dip in the average sales price of agricultural land. Instead, figure 2 suggests that the average farmland sale prices stayed fairly constant at around $5000/acre over the 2000s decade, which was in part due to the growing significance of agricultural markets, exemplified by the surging biofuels market (Wallander, et al. 2011) and rising demand for U.S. grain exports (Gloy, et al. 2011).

To further explore these issues especially the change in the effect of proximity to ethanol plants, I first estimate two hedonic models as the benchmark model shown in table 9, which can be viewed as the reduced-form estimations for the instrumental variables approach. Specifically, model (I) uses all observations without matching while model (II) uses a matched sample based on proximity to ethanol plants but does not control for the potential endogeneity of plant location using instrumental variables. The hedonic estimate on the variable *Dist_Ethanol * Post construction dummy* shows that proximity to

Nominal farmland values ($/acre)	(I)		(II)	
	Coef.	Robust SE	Coef.	Robust SE
Distance to nearest ethanol plant	-0.87	5.77	4.07	13.08
Dist_Ethanol * Post construction dummy	-5.86	3.93	-13.35	9.79
Assessed land value % of total assessed	-3771.54***	145.03	-3354.51***	374.90
Total acres	-26.07***	1.04	-40.15***	3.07
Total acres squared	0.013***	0.0017	0.08***	0.01
NCCPI	0.0028	0.026	0.037	0.05
Prime farmland	-71.25	116.24	-258.21	266.94
Steep slope (>15 degrees)	-100.49*	58.51	309.05	255.31
Building area % of parcel	52.97	268.36	-550.92	394.57
Forest area % of parcel	14.17	159.23	-403.03	521.55
Wetland area % of parcel	-112.28	876.67	845.14	3332.38
Distance to highway ramp	-37.23**	15.18	-30.05	27.94
Distance to nearest city	-62.54***	8.01	-25.52	15.68
Incremental distance to 2^{nd} nearest city	-36.40***	5.99	-31.04**	13.68
Surrounding population within 25 miles	0.62**	0.31	-0.36	0.80
Gravity index of three nearest cities	2.87E-04*	0.0002	0.77	0.56
Distance to railways	-3.69	17.35	-0.28	32.24
Distance to nearest grain elevator	2.63	9.89	-35.76	26.54
Distance to nearest agricultural terminal	-33.15***	5.41	-4.50	11.77
Intercept	14456.89***	848.59	15742.13***	3340.06
County FE	Yes		Yes	
Year FE	Yes		Yes	
Adjusted R^2	0.2616		0.2409	
Number of observations	16434		3443	

Continued

Table 9. Hedonic Regressions with Structural Changes of Proximity to Ethanol Plants

Table 9 continued

Note: *, **, and *** indicates the coefficient is significant at 10%, 5% and 1% level, respectively. 50 county fixed effects are included in the model.

ethanol plants became a positive influence after construction of these plants; however, this positive capitalization effect is not statistically significant. Most of the estimates of other variables were intuitive: bad soil quality or presence of steep slope decreased farmland values, while proximity to urban areas or highway ramps led to an increase. The significant coefficient on acres squared implied a nonlinear relationship between the per-acre farmland values and total acreage. Model (II) uses the matched sample which controls for the differences in observable characteristics between parcels closer to ethanol plants and those farther away, and this leads to the insignificance of some variables such as proximity to cities and agricultural terminals. These results are preliminary since they did not control for the potential endogeneity of the location of ethanol plants. Nonetheless, the results are suggestive and provide ample motivation to further investigate potential structural change in these effects using instrumental variables estimation with matched samples.

Table 10 shows the comparison of the difference-in-means of the covariates between treatment and control groups for the raw sample before matching and the matched sample after PSM. The naive control group for the raw sample was constructed as if they were matched using the same cutoff values in timing or distance. To make the second step matching work, the 50 county fixed effects were replaced by six crop reporting district dummies because of too few observations within each county. The distance cutoffs used to construct matched samples based on proximity to agricultural market channels are 10 miles, 5 miles, and 15 miles for ethanol plants, grain elevators, and agricultural terminals, respectively. Table 10 reveals that without matching, there were systematic distributional differences between the naive control group and the treatment group, which, as a result, led to biased estimates in the standard hedonic approach. In contrast, at the cost of reduced sample size, these differences were successfully removed through PSM, which assures that conditioning on estimated propensity score, there is no remaining distributional differences left for the covariates between the treatment and control groups (Lynch, et al. 2007). In a word, table 10 illustrates the necessity and advantages of propensity score matching in addressing the sample selection bias inherent in the standard hedonic method. Figure 11 in the Appendix A shows that the estimated propensity score for the treatment and control groups overlap with each other.

Table 11 presents the main results of regressions on matched samples for each of the three types of agricultural market channels. In particular, instrumental variables regression is used to estimate the effects of proximity to ethanol plants, while hedonic regression is used for grain elevators and agricultural output terminals. The main result provides evidence for positive and significant marginal value of being within close proximity to an ethanol plant following construction of seven ethanol plants in or near western Ohio in late 2006-early 2007.

Covariates	Matching on timing: month of ethanol plant construction		Matching on distance					
			Ethanol plant		Grain elevator		Agricultural terminal	
	Unmatched	Matched	Unmatched	Matched	Unmatched	Matched	Unmatched	Matched
Sale year			-0.7***	-0.1	2005.4	0	2005.4	0
Assessed land value %	-0.0426***	-0.0042	-0.156***	-0.0100	0.7631	0.0066	0.7243	-0.0115
Total acres	-5.8770***	0.9380	-4.953***	-1.7080	46.8740	1.0190	50.0400	1.0580
Total acres squared	662.30***	490.9000	534.30	-406.70	4745.40	120.10	6790.00	1484.60
Agricultural Productivity Variables								
NCCPI	-121.60***	-5.6000	369.40***	-52.50	5870.70	-47.60	5717.80	-3.70
Soil class 2 area % of parcel	0.0110*	-0.0039						
Soil class 4 area % of parcel	-0.0058	0.0043	0.0422***	-0.0417***	0.1283	0.0059	0.1702	0.0056
Steep slope (>15 degrees)	0.0483***	0.0025	0.1688***	0.0005	0.0824	-0.0019	0.1361	0.0313**
Urban Influence Variables								
Total urban popu within 25 miles	20.9200***	-3.1600	97.7900***	-2.6	246.47	8.9700**	242.27	-6.69
Distance to highway ramp	-0.2573***	0.0063	0.0630	0.0440	4.9915	0.0535	6.2736	-0.1780
Distance to nearest city	-4.495***	0.0900	-13.885***	-0.3900	44.6890	-0.5520	48.2500	-0.2860

Continued

Table 10. Difference in Means of the Covariates between Treatment and Control Groups for the Raw and Matched Samples

Table 10 continued

Agricultural Market Influence Variables								
Distance to nearest ethanol plant	5.9380***	-0.0160			35.8380	-0.1260	38.5880	1.2040*
Distance to nearest grain elevator	0.2630***	-0.1070	6.0960***	-0.3842*			11.2920	0.6140**
Distance to nearest agricultural terminal	-1.4780***	-0.0380	2.8640***	-1.2120*	51.9860	0.3640		
Environmental Influence Variables								
Forest area % of parcel	0.0434***	-0.0002	0.0890***	-0.0033	0.0770	-0.0009	0.0851	0.0069
Wetland area % of parcel	0.0010*	-0.0001	0.0014***	0.0001	0.0039	-0.0009	0.0023	-0.0003
Location Fixed Effects								
County FE	Yes	Yes						
Northwest crop reporting district FE			0.0886***	0.0154	0.3336	0.0023	0.2781	-0.0228
Central crop reporting district FE			0.0615***	-0.0144	0.2158	-0.0170*	0.3851	0.0160
West central crop reporting district FE			-0.3096***	-0.0011	0.4017	0.0124	0.3369	0.0068
Observation	12969	9880	9880	2082	9880	8055	6702	4174

Note: *, **, and *** indicates the difference in means of the covariates between treatment and control groups is significant at 10%, 5% and 1% level, respectively. In the first step of matching on timing, a parcel is considered to be in the treatment group if it is sold after the construction of nearest ethanol plant. The matched sample forms the full sample for the second step of matching on proximity to agricultural markets. A parcel is considered to be in the treatment group if it is located within 10 miles, 5 miles, and 15 miles for ethanol plants, grain elevators, and agricultural terminals, respectively.

Specifically, results from the IV estimation with the matched sample suggest that the marginal value of farmland increases by $46 per mile per acre within proximity to the nearest ethanol plant following construction of these plants. By comparison, the effect of proximity to nearest city center and second nearest city is $30-66 and $30-40 per mile per acre, respectively. In contrast, there is no declining farmland price gradient over distance to ethanol plants. Since there are no ethanol plants before late 2006, this coefficient on the interaction term can be interpreted as the positive effects of proximity to newly constructed ethanol plants.

To test the validity and relevance of my instruments, I run a series of tests and robustness checks. Table 23 in the Appendix A presents results of the first stage regressions of the potentially endogenous variables. The significant and negative coefficient of the proposed instrument *capacity-weighted distance to other ethanol plants* in table 23 column (II) confirms my conjecture of spatial competition among ethanol plants. Table 25 shows the test statistics for the weak identification of the instruments as well as the test of overidentifying restrictions. I cannot reject the overidentification test based on the Hansen J statistic (Hansen 1982) , acknowledging the strong assumption of at least one valid instrument. Table 24 further shows an example of a regression of instruments on other exogenous covariates, which reveals that the instrument is not correlated with most covariates and serves as an indirect test for the validity of instrument variables. In addition, the Kleibergen-Paap Wald F statistic reveals that these instruments are relevant and not weak (Kleibergen and Paap 2006; Stock and Yogo 2005). A comparison between table 2 and table 11 column (a) reveals that, regardless of the significance, the standard

Nominal farmland values ($/acre)	(a) Ethanol Plant		(b) Grain Elevator		(c) AG Terminal	
	Coef.	SE	Coef.	SE	Coef.	SE
Distance to nearest ethanol plant	16.53	49.59	5.34	7.95	-2.49	10.75
Dist_Ethanol * Post construction	-46.39**	19.57				
Assessed land value % of total assessed	-3331***	418.53	-3654 ***	221.23	-3142***	259.13
Total acres	-36.8***	2.92	-35.17***	1.81	-50.44***	2.40
Total acres squared	0.06***	0.01	0.038***	0.005	0.093***	0.0159
NCCPI	-4.76E-05	0.05	0.011	0.034	-0.0153	0.0488
Prime farmland	-358.02	289.08	-212.71	181.27	-336.86	237.51
Steep slope (>15 degrees)	276.95	333.64	-123.26	92.75	-10.12	116.38
Building area % of parcel	-345.72	408.32	-541.92	400.67	-543.18	461.43
Forest area % of parcel	-35.50	646.12	94.29	275.95	-315.08	348.59
Wetland area % of parcel	-997.00	3621.69	-265.83	879.32	1092.43	1913.35
Distance to highway ramp	-24.40	30.66	-28.75	21.53	-74.25***	26.31
Distance to nearest city	-30.86**	15.66	-28.91***	11.28	-66.17***	16.89
Incremental distance to 2nd nearest city	-39.59***	14.34	-12.02	8.17	-33.83***	11.99
Surrounding population within 25 miles	-0.74	0.77	0.73	0.46	1.307**	0.556
Gravity index of three nearest cities	1.24*	0.67	-0.05	0.06	0.0214	0.076
Distance to railways	20.34	35.29	-15.35	23.6	-13.77	57.56
Distance to nearest grain elevator	-27.70	31.91	5.36	20.93	33.35*	18.95
Dist_Grain * Post construction dummy			-57.50**	29.47		
Distance to nearest agricultural terminal	6.52	16.19	-15.72*	8.22	-48.48***	20.09
Dist_Terminal * Post construction					18.12**	7.86
Intercept	10261.9***	1858.68	9721.26***	744.78	23762***	1966.72
Year FE			yes			
Adjusted R^2	0.2473		0.2619		0.2885	
Number of observations	3443		8123		4864	

Continued

Table 11. Structural Change in the Effects of Proximity to Agricultural Markets Channels – Regressions on the Matched Sample

Table 11 continued

Note: *, **, and *** indicates the coefficient is significant at 10%, 5% and 1% level, respectively. 50 county fixed effects are included in the model. I use an instrumental variables regressions for ethanol plants and hedonic regressions for grain elevator and agricultural output terminals.

hedonic estimates yield similar signs with the instrumental variables approach. This is because the suspected endogenous variable *distance to nearest ethanol plant* is only slightly endogenous (with a p-value of 0.064), according to the endogeneity test shown in table 25 panel (III) (Baum, et al. 2003). However, this comparison between the standard hedonic estimates and the instrumental variables estimates confirms the endogeneity of proximity to ethanol plants, which, if left uncontrolled for, would result in a downward bias in the standard hedonic estimates due to unobserved characteristics. This underestimation might result from unobserved characteristics that capture rural remoteness or other undesirable traits.

Table 26 in the Appendix A presents another set of robustness checks for the validity of the instruments by regressing farmland values directly on instruments and other exogenous covariates as opposed to endogenous variables. Table 26 panel (I) uses all observations before matching while panel (II) and (III) use the matched sample. The results show that *capacity-weighted distance to other agricultural terminals* after construction could be slightly endogeneous in the raw sample before matching, but this problem is eliminated after matching. Table 26 panels (II) and (III) show that *capacity-*

		(a) Ethanol Plant		(b) Grain Elevator		(c) AG Terminal	
		Coef.	SE	Coef.	SE	Coef.	SE
Panel I: 1 to 2 nearest neighbor matching	Dist_Ag Market	48.77	68.52	32.00	27.26	-52.16***	12.75
	Dist_Ag Mkt * Post_Dummy	-62.15***	20.57	-121.82**	37.66	20.63**	9.22
	Number of observations	2721		5264		3658	
	Adjusted R^2	0.2505		0.2675		0.2955	
Panel II: 1 to 1 nearest neighbor matching	Dist_Ag Market	173.43	120.91	5.36	20.93	-44.56***	16.14
	Dist_Ag Mkt * Post_Dummy	-66.85***	25.12	-57.50*	29.47	21.32**	10.85
	Number of observations	2151		8123		2828	
	Adjusted R^2	0.2023		0.2619		0.3058	
Panel III: Mahalanobis metric covariate matching	Dist_Ag Market	77.25	110.33	32.96	25.55	-20.20*	12.75
	Dist_Ag Mkt * Post_Dummy	-84.34*	45.12	-74.41**	31.90	16.74	14.25
	Number of observations	2175		7338		2597	
	Adjusted R^2	0.2238		0.2762		0.3025	
Panel IV: Kernel matching	Dist_Ag Market	49.13	38.03	3.49	14.35	-43.05***	8.21
	Dist_Ag Mkt * Post_Dummy	-53.84***	15.92	-30.24*	17.10	21.27***	6.08
	Number of observations	11440		15398		7759	
	Adjusted R^2	0.2654		0.2715		0.2943	

Continued

Table 12. Robustness Checks of Alternative Matching Algorithms

Table 12 continued

Note: *, **, and *** indicates the coefficient is significant at 10%, 5% and 1% level, respectively. 50 county fixed effects are included in the model. The distance cutoffs used to construct matched samples based on proximity to agricultural market channels are 10 miles, 5 miles, and 15 miles for ethanol plants, grain elevators, and agricultural terminals, respectively.

weighted distance to other ethanol plants could still be endogenous after the construction of ethanol plants, and it may present as a better alternative to just use *capacity-weighted distance to other agricultural terminals* as the instruments for the matched sample. Columns (b) and (c) in table 11 also reveal a stronger influence of proximity to grain elevators as well as a reduction in the magnitude and significance of the effect of proximity to agricultural terminals after early 2007 due to competition from the newly constructed ethanol plants. Specifically, proximity to grain elevators did not exert significant influence in surrounding farmland values before 2006, but became a positive and significant determinant after the ethanol market expansion in Ohio. This result is intuitive because local grain elevators meet part of the increased demand for corn due to construction of ethanol plants. In addition, I find that the marginal value of being close to an agricultural terminal reduces from $48 to $30 per mile per acre after early 2007, which suggests that the newly constructed ethanol plants constitute a significant competing source of demand for grains for traditional agricultural output terminals (Nickerson, et al. 2012). This also lends support for my instrumental variables approach which relies on the spatial competition among ethanol plants and agricultural terminals.

		(a) Ethanol Plant		(b) Grain Elevator		(c) AG Terminal	
		Coef.	SE	Coef.	SE	Coef.	SE
Panel I: Reduced distance cutoff for proximity	Dist_Ag Market	402.67	317.67	4.82	22.53	-68.19***	22.42
	Dist_Ag Mkt * Post_Dummy	-64.85*	36.50	-63.58**	30.83	45.10***	13.73
	Number of observations	1409		6940		1917	
	Adjusted R^2	0.2551		0.2659		0.2886	
Panel II: Alternative reduced distance cutoff for proximity	Dist_Ag Market	57.81	51.02	22.39	17.91	-53.04***	15.52
	Dist_Ag Mkt * Post_Dummy	-60.08***	20.59	-60.99**	24.30	24.82**	10.26
	Number of observations	2505		9455		3104	
	Adjusted R^2	0.2169		0.2694		0.2867	
Panel III: Increased distance cutoff for proximity	Dist_Ag Market	47.20	39.92	4.13	14.96	-40.14***	8.83
	Dist_Ag Mkt * Post_Dummy	-59.89***	16.74	-37.87**	19.24	14.10**	6.62
	Number of observations	4390		11698		6492	
	Adjusted R^2	0.2303		0.2714		0.2843	
Panel IV: Change timing 6 months earlier	Dist_Ag Market	22.68	44.07	21.21	18.69	-42.84***	8.87
	Dist_Ag Mkt * Post_Dummy	-39.92**	16.20	-49.43**	23.19	18.07***	6.27
	Number of observations	3443		9455		6492	
	Adjusted R^2	0.2395		0.2693		0.2847	
Panel V: Change timing 1 year earlier	Dist_Ag Market	-4.69	44.17	9.89	18.90	-43.51***	8.93
	Dist_Ag Mkt * Post_Dummy	-8.91	14.60	-23.11	23.03	17.43***	6.09
	Number of observations	3443		9445		6492	
	Adjusted R^2	0.2403		0.2403		0.2847	

Continued

Table 13. Robustness Checks using Alternative Distance and Timing Cutoffs

Table 13 continued

Panel VI: Change timing from construction to plant opening	Dist_Ag Market	14.26	24.35	4.33	14.66	-36.85***	8.60
	Dist_Ag Mkt * Post_Dummy	-28.94*	15.87	-67.21**	22.15	9.20*	4.86
	Number of observations	3443		10880		6492	
	Adjusted R^2	0.2502		0.2704		0.2847	
Panel VII: Log-linear specification	Dist_Ag Market	0.0014	0.0088	0.0006	0.0030	0.0099***	0.0020
	Dist_Ag Mkt * Post_Dummy	-0.0062*	0.0035	-0.0088**	0.0040	-0.0042***	0.0015
	Number of observations	3443		10879		4864	
	Adjusted R^2	0.2276		0.325		0.4005	

Note: *, **, and *** indicates the coefficient is significant at 10%, 5% and 1% level, respectively. 50 county fixed effects are included in the model. The distance cutoffs used to construct matched samples based on proximity to agricultural market channels in panel I are 6 miles, 3 miles, and 10 miles for ethanol plants, grain elevators, and agricultural terminals, respectively. In panel II the distance cutoffs are 8 miles, 4 miles, and 12 miles for ethanol plants, grain elevators, and agricultural terminals, respectively. In panel III the distance cutoffs are 12 miles, 6 miles, and 18 miles for ethanol plants, grain elevators, and agricultural terminals, respectively.

newly constructed ethanol plants. Table 13, on the other hand, tests the robustness of the results by altering the cutoff distances used to define the spatial proximity and the timing used to define when the effect of ethanol plants start to kick in. Results reveal that there is evidence of expectations before the construction of ethanol plants; however, the expectations argument is only relevant 6 months before the plant construction. The log-linear model specification also reveals a similar conclusion. These robustness checks

		Coef.	Std. Err.
Panel I: Exclude 2 nearest ethanol plants and agricultural terminals	Dist_Ag Market	44.14	35.48
	Dist_Ag Mkt * Post_Dummy	-48.68***	17.46
	Number of observations	3541	
	Adjusted R^2	0.243	
Panel II: Exclude 3 nearest ethanol plants and agricultural terminals	Dist_Ag Market	43.81	36.93
	Dist_Ag Mkt * Post_Dummy	-53.08***	19.52
	Number of observations	3541	
	Adjusted R^2	0.2426	
Panel III: Just include agricultural terminals	Dist_Ag Market	34.59	30.11
	Dist_Ag Mkt * Post_Dummy	-32.47*	17.09
	Number of observations	3541	
	Adjusted R^2	0.2377	
Panel IV: Only use parcels within the Corn Belt boundary	Dist_Ag Market	23.61	51.68
	Dist_Ag Mkt * Post_Dummy	-49.23***	22.7
	Number of observations	3253	
	Adjusted R^2	0.2465	
Panel V: No other covariates except the instruments	Dist_Ag Market	27.61	40.54
	Dist_Ag Mkt * Post_Dummy	-40.25**	20.54
	Number of observations	3430	
	Adjusted R^2	0.0996	

Table 14. Robustness Checks using Alternative Definitions of Instruments
Note: *, **, and *** indicates the coefficient is significant at 10%, 5% and 1% level, respectively.

I test the stableness of the results by employing alternative ways to construct the matched sample and alternative model specifications. Table 12 presents results using alternative matching algorithms, including propensity score matching using one or two nearest neighbors instead of four, as well as completely different matching estimators which includes covariate matching (Rubin 1980) and kernel-based matching (Heckman, et al. 1998). Despite greater magnitude, the robustness checks show a similar conclusion as the main specification – a significant and positive effect of proximity to newly constructed ethanol plants after their construction. Results also confirm that finding regarding the reduced impact of proximity to agricultural output terminal due to strong competition of newly constructed ethanol plants. Table 13, on the other hand, tests the robustness of the results by altering the cutoff distances used to define the spatial proximity and the timing used to define when the effect of ethanol plants start to kick in. Results reveal that there is evidence of expectations before the construction of ethanol plants; however, the expectations argument is only relevant 6 months before the plant construction. The log-linear model specification also reveals a similar conclusion. These robustness checks indicate that my results are stable across different model specifications and matching algorithms.

Table 14 further tests for the validity of instruments and the assumption of the hedonic market. The instruments relying on the distances to other plants would not be valid if more than one agricultural market could affect surrounding farmland values. I test that by excluding two or three nearest ethanol plants and agricultural output terminals in the construction of instruments, which suggests that only other agricultural markets used in the instruments are farther away to affect the farmland values for parcels closer to a particular ethanol plant. These rests on the assumption that agricultural markets sufficiently far away from the farmland parcel would not have a direct impact on its sale price. Table 14 panels I, II and III reveal similar results as in table 10, suggesting that my instruments are valid and unlikely to be endogenous. In particular, panel III just uses agricultural terminals when constructing the instruments, and the similarity with the main specification echoes with the robustness check shown in table 26 panel (III) in the Appendix A, suggesting that overall the instrumental variables regressions yield robust and reliable estimates. Using estimates from semiparametric regressions, Figure 12 in the Appendix A shows that the effect of proximity to ethanol plants is likely localized within 15-20 miles, further validating the validity of the instruments. Table 14 panel IV uses only parcels from counties with at least 5 million bushels of corn production in 2010 or within the Corn Belt boundary to ensure the farmland sales data used in this study could be considered as part of one single farmland market. Table 14 panel V only keeps instruments in the regressors and drops other covariates. Both panels IV and V yield qualitatively similar results as the main specification.

Conclusion

The first decade of the 2000's saw dramatic changes in the forces that influence farmland values. On one hand, rapid expansion of biofuels markets supported by federal energy policies has dramatically increased demand for corn, which elevated agricultural commodity prices and farmland values (Wallander, et al. 2011). On the other hand, the residential housing market bust in 2006 that precipitated the Great Recession had a substantial negative effect on the value of exurban farmland proximate to urban areas. Using a dataset of parcel-level farmland sales in western Ohio from 2001 to 2010 and a quasi-experimental design, this study tests the common wisdom that the rise of ethanol industry helped the farm sector withstand this downturn (Nickerson, et al. 2012). The identification strategy relies on the observed opening of seven ethanol plants in or near western Ohio and the competitive pressures they face in the spatial competition with other ethanol plants and agricultural output terminals. Two instruments are constructed based on this idea of spatial competition among agricultural markets to control for the potentially endogeneous proximity of farmland parcels to nearest ethanol plants. Propensity score matching is also used to control for potential sample selection bias resulting from systematic differences in observable parcel characteristics between parcels closer to agricultural market channels versus those farther away.

The main results from the instrumental variables estimation with the matched sample suggest that the marginal value of farmland increases by $46 per mile per acre with close proximity to the nearest ethanol plant following construction of these plants. Despite a short-term effect with data only for the 2000s, these results demonstrate the growing importance of the biofuels market for farmland values and show that proximity to ethanol plants is becoming a more significant determinant of agricultural land values. This study suggests the need to explicitly incorporate agricultural market influence variables such as proximity to ethanol plants and other agricultural market channels in modeling the determinants of U.S. farmland values. This study also confirms the endogeneity concerns of the proximity to ethanol plants and reveals systematic differences in observable characteristics between parcels closer to an agricultural market channel and those farther away, which could lead to biased estimates if left uncontrolled for as in the standard hedonic price method. The quasi-experimental design that combines matching and instrumental variables approach employed here presents a superior alternative.

Ethanol is now a critical part in the corn industry supply chain and the year of 2010 marks the first time that corn usage for ethanol production exceeds usage for feed stock (Wallander, et al. 2011). However, until recently, ethanol development and utilization have been largely dependent upon government subsidies and other policy support. There is ongoing debate regarding the welfare impacts of ethanol policy and resulting ethanol market expansion, including its impacts on farmer income, commodity prices, farmland values, greenhouse gases, and energy portfolio (Cappiello and Apuzzo 2013; Rajagopal, et al. 2011; Tiffany 2009). In addition to many criticisms of various subsidies offered to ethanol producers, rising concerns are raised regarding the environmental quality impacts of ethanol policy through its incentives for corn expansion (Cappiello and Apuzzo 2013). This study engages in this debate by providing a piece of evidence on the capitalization effect of proximity to ethanol plants. With many subsidies already or slated to be terminated, it poses an intriguing policy question that how these downward pressure on ethanol development would affect the welfare effects of the ethanol market expansion, and in particular its capitalization in commodity prices and farmland values.

Chapter 3: Alternative Nutrient Management Policies and the Trade-offs between Agricultural Profits and Water Quality Improvements

Introduction

Excessive nutrient runoff from agricultural production contributes to freshwater eutrophication and coastal hypoxia across the United States and globally. In 2011, a harmful algal bloom of unprecedented size and severity occurred in the western basin of Lake Erie (Michalak, et al. 2013). The Lake Erie watershed is the most populous of all the Great Lakes and such events pose significant risks to ecosystem services provided by this vital lake, including recreation opportunities, public health and safe drinking water. To address the agricultural nutrient loss problem, substantial efforts have been made at the federal level to promote adoption of best management practices through voluntary payments for conservation programs. In 2010, the national total payments for conservation program amounted to $3.5 billion, or around $10 per acre of farmland. Despite these efforts, no measurable improvements in downstream water quality has been found, as evidenced by the record-setting HAB in Lake Erie in 2011 and the a two-day shutdown of the City of Toledo’s public water system last summer. The stark contrast suggests that the current voluntary-based policy may not be sufficient to improve water quality. This naturally leads to the question of whether more stringent agri-environmental policies can improve the cost-effectiveness of agricultural pollution control, and what are the trade-off between agricultural profits and ecosystem benefits.

Evaluation of agri-environmental policies is complicated by multiple sources of heterogeneity: farmers' responses to market-based incentives or regulation will vary depending on their own behavioral motivations, characteristics of their land, and locational factors, all of which are heterogeneous. Many studies have shown that farmers' socioeconomic and demographic characteristics are important in driving the adoption of conservation practices (Featherstone and Goodwin 1993; Norris and Batie 1987), and behavioral preference heterogeneity, including environmental stewardship and perceived efficacy of the policy, are important factors in explaining the diversity of responses of farmers to incentive-based programs (Howard and Roe 2013; Wilson, et al. 2014). However, mainly due to lack of data, individual field-level operator characteristics have been largely ignored in aggregate structural models on agricultural crop and input demand choices (Denbaly and Vroomen 1993; Griliches 1958; Moschini 1988; Shumway and Alexander 1988), in fertilizer demand models that examine heterogeneity in demand elasticity(Hansen 2004; Hansen and Hansen 2014), in economic models of conservation practices and nutrient management decisions (Kurkalova, et al. 2006; Laukkanen and Nauges 2014; Wu, et al. 2004), in economic models of land conservation and spatial targeting (Ando and Mallory 2012; Babcock, et al. 1997), and in recent integrated ecological-economic models (Hendricks, et al. 2014; Rabotyagov, et al. 2014). Williamson (2011) is one recent study on nitrogen fertilizer demand that incorporates farm-level land and farmer characteristics using USDA's Agricultural Resource

Management Survey in 2001 and 2005; however, it is reduced form in nature and thus not well-suited for policy analysis. Timmins and Schlenker (2009) and others have argued that a structural model is needed for the welfare analysis of non-marginal policy changes. A structural model of fertilizer demand is thus needed when there are yield setbacks from fertilizer reduction or substitution among different nutrient inputs, or when analyzing a non-marginal policy change like a 100% fertilizer tax. In addition, modeling policy scenarios and quantifying its welfare effects and trade-offs, which often requires a structural model, provide useful information for policymakers and the public. In sum, as far as I know, there lacks an empirically-based ecological-economic model of farmer decision making with a structural foundation, and one that explicitly accounts for both spatial land heterogeneity as well as farmers' behavioral heterogeneity.

The objective of this study is to quantify the social welfare impacts of alternative nutrient management policies in terms of both farmer profits and ecosystem damages, and examine whether and how heterogeneity in fertilizer demand elasticities impact the welfare analysis. I hypothesize that the optimal nutrient management policy varies depending on how farmer heterogeneity and spatial heterogeneity interact with each other in influencing farmers' land use land management decisions. For example, if the derived elasticity of fertilizer demand decreases with land quality, then a fertilizer tax that targets fields with top soil quality and leads to a higher proportion of reduction in fertilizer application may be optimal. On the other hand, farmers with similar preferences for environmental stewardship or other behavioral attributes may be spatially correlated—for example, due to common media coverage of local agri-environmental problems. In this case, a zonal tax based on proximity of certain sub-watersheds to the downstream distressed ecosystem might be warranted.

To quantify the social welfare impacts of alternative agri-environmental policies, I develop a field-level structural econometric model that, in a profit-maximizing framework, simultaneously models crop choice, fertilization frequency choice and nutrient management decisions. In a simultaneous equations framework, my model improves on the previous reduced-form models by explicitly accounting for the selectivity of crop and nutrient management choices and the joint nature of all input demand decisions. Moreover, following the suggestion of Timmins and Schlenker (2009), I first estimate the key parameter of mean derived elasticity of phosphorus fertilizer demand using a reduced form panel data model and then combine it with the structural estimation. This structural approach allows us to quantify the social welfare impacts of different agri-environmental policies by quantifying the changes in farmer profits and ecological damages resulting from a change in phosphorus application rates. In particular, I evaluate the environmental and economic trade-offs from alternative policies like uniform fertilizer tax, spatially targeted tax based on land characteristics or location of the parcel, as well as an educational campaign to raise awareness about nutrient stewardship and a limit on fertilizer rate.

I apply this model to the Maumee River Watershed, which is the largest in the Great Lakes watersheds and contributing by far the largest volume of sediment and nutrient loadings into Lake Erie, which contributes to excessive, harmful algal blooms (HABs) and other water quality problems in Lake Erie (Reutter, et al. 2011). Data on farmers' economic, behavioral and land characteristics come from a comprehensive survey of 7,500 farmers in this watershed with extensive questions on farmers' field level choices of multiple land management practices in 2013, demographics and risk attitudes of farmers, and farm and field level spatial characteristics. Scientific studies have suggested that at least 40% reduction in phosphorus runoff from Maumee is likely needed to significantly reduce the incidence and intensity of HABs in Lake Erie (Ohio Environmental Protection Agency 2013).

The main results reveal that farmers respond differently to a phosphorus fertilizer price change depending on their environmental attitudes, quality of land, and their crop and fertilization frequency choices. The estimated price elasticity of farmers' phosphorus fertilizer demand is relatively inelastic, ranging from -0.26 to -0.60. Results show that neither a phosphorus fertilizer tax across the watershed nor an educational campaign on environmental stewardship would be sufficient to achieve the policy goal of 40% reduction in phosphorus runoff from Maumee. I also find that spatial targeting, such as phosphorus tax targeted towards ecologically sensitive subbasins, improves the cost-effectiveness of agri-environmental policies when only costs to farmers are considered; while a simpler policy such as a 50% uniform phosphorus tax would outperform other alternatives when the cost-effectiveness is measured as phosphorus reduction given net policy costs from an overall social welfare perspective.

This study makes several contributions to the literature on the modeling of agro-ecosystems and agri-environmental policies. First, I empirically estimate a structural econometric model of crop and nutrient management choices that is better suited for the welfare analysis of policies, especially for non-marginal policies such as a 100% phosphorus tax, than previous numerical simulation or reduced-form models. Second, the structural model accounts for farmers' heterogeneity in phosphorus fertilizer demand using a unique and rich dataset on field, farm, and farmer characteristics from a survey of 1,551 farmer respondents from western Lake Erie basin. Third, the estimation makes use of two additional hypothetical questions on phosphorus applications and shows that the identification of a key parameter - mean elasticity of phosphorus fertilizer demand - using a reduced-form panel data analysis complements the structural model estimation. Finally, the analysis has strong policy implications for the sustainability of Lake Erie agri-ecosystem and regions alike, and the key results show that neither an education campaign on nutrient stewardship nor a 100% fertilizer tax would achieve the policy goal of 40% reduction in phosphorus loadings needed to alleviate the harmful algal blooms in Lake Erie.

Literature Review on Fertilizer Demand and Agri-Environmental Policies

There has been a vast literature on the analysis of agricultural input demand and fertilizer demand in particular. Most of the studies dating back to Griliches (1958), however, focus on the aggregate fertilizer demand functions and their dynamics over time using aggregate, annual time-series data at the national or regional level (Binswanger 1974; Carman 1979; Denbaly and Vroomen 1993; Gunjal, et al. 1980; Heady and Yeh 1959; Penm and Vincent 1987). Most of these studies find that the price elasticity of nitrogen and phosphorus fertilizer demand to be inelastic, ranging from -0.20 to -0.95. For example, Denbaly and Vroomen (1993) use co-integrated and error-corrected models with U.S. time series data from 1964 to 1989 and report a short-run Marshallian elasticity of -0.25 and a long-run elasticity of -0.37 for corn producers' phosphorus fertilizer demand. Of the early studies of fertilizer demand, Pitt (1983) is a rare application that uses data at individual farm level, but no variables that capture the land heterogeneity or farmers' demographic and socioeconomic characteristics are used in the estimation. In a simultaneous equations framework, Laukkanen and Nauges (2014) use a national sample of Finnish grain farms from 1996 to 2005, and quantify the effects of European Union's agri-environmental payments on farmers' decisions and on fertilizer use. While their study focuses on the elasticities of land and fertilizer use with respect to grain or environmental subsidies, they also report the price elasticity of fertilizer demand to be -0.91.

More recently, Williamson (2011) and Ricker-Gilbert, et al. (2011) are two examples that incorporate the increasingly available farm-level data on land and farmer characteristics into the model of fertilizer demand. Williamson (2011) finds that the estimated price elasticity of nitrogen demand ranges from -1.67 to -1.87 using farm-level microdata from USDA's Agricultural Resource Management Survey (ARMS) in 2001 and 2005; while Ricker-Gilbert, et al. (2011) focus on the crowding out effect of fertilizer subsidy on commercial fertilizer use in Malawi. However, neither study focuses on the heterogeneity in the price responsiveness of fertilizer demand. Hansen (2004) uses a micro panel data of 1350 Danish crop farms from 1983 to 1991, and finds that the mean nitrogen fertilizer demand elasticity is -0.45. More importantly, Hansen (2004) introduces heterogeneity in demand elasticities by letting the price coefficient to be a quadratic function of profit shares of nitrogen fertilizer, and he finds a significant standard deviation in the estimated elasticity of 0.24. However, due to lack of data on land and farmer characteristics, Hansen (2004) just reveals or shows heterogeneity among farms but does not explain the sources of heterogeneity.

While heterogeneity in demand is not the focus of most empirical studies of fertilizer demand, it is a frequent topic in related literature such as demand for food, alcohol, cigarette or gasoline. For example, Ayyagari, et al. (2013) employ a latent class model to model heterogeneity in price elasticities in demand for alcohol; Jacobsen (2013) allows the parameters in demand elasticity for gasoline to vary by income, education, city size and race. Meier, et al. (2010) account for heterogeneity in price sensitivity in demand for alcohol by estimating the model separately by population subgroups by age, gender, or level of drinking. Harding and Lovenheim (2014) incorporate the demographic characteristics in estimating heterogeneous demand for nutrition in a quadratic Almost Ideal Demand System model. Both Gillingham (2014) and Liu (2014) introduce heterogeneity in demand elasticity of gasoline by interacting demographic or state-level socioeconomic covariates with gasoline prices, and Gillingham (2014) also employs a quantile regression and k-means cluster analysis to uncover the heterogeneity.

The strand of literature that does examine the heterogeneity in fertilizer demand, which often results from spatial differences in land quality, or spatially differentiated agricultural pollution control policies, is mostly theoretical models with numerical simulations (Claassen and Horan 2001; Goetz and Zilberman 2000; Iho and Laukkanen

2012; Lankoski, et al. 2010; Lankoski and Ollikainen 2003; Xabadia, et al. 2008). For example, Goetz and Zilberman (2000) develop a theoretical model for the socially optimal management of phosphorus runoff taking into account of dynamics over time and spatial heterogeneity represented by land quality. In a parametric analytical model, Lankoski, et al. (2010) represent both the agricultural profits and ecosystem damages as function of fertilizer application rates, and quantify numerically the impacts of agri-environmental policies, including spatially-differentiated fertilizer tax, on farm income, nitrogen runoff damage as well as the enforcement costs. Another related literature on land and biodiversity conservation also illustrates the potential efficiency gains from spatially-targeted policy instruments as opposed to uniform policies (Ando and Mallory 2012; Babcock, et al. 1997; Newburn, et al. 2006). Most of these recommendations stem from the recognition of heterogeneous environmental impacts or benefits resulting from spatially-differentiated landscapes, as opposed to heterogeneous preferences or decision-making process among landowners.

On the other hand, the literature on the adoption of best management practices (BMP) have always focused on the importance of field-level land and operator characteristics by making use of farmer survey data (Featherstone and Goodwin 1993; Norris and Batie 1987). Some recent studies have suggested that important sources of preference heterogeneity, including environmental stewardship and perceived efficacy of the policy, are important factors in explaining the diversity of responses of farmers to incentive-based programs (Howard and Roe 2013; Wilson, et al. 2014). However, typically these models are not spatially articulated and thus cannot be used to assess the environmental and welfare implications of policies encouraging BMP adoptions.

There is also a growing literature on integrated ecological-economic models that link models of agricultural land use land management with analytical or biophysical models of agricultural nutrient flows such as Soil and Water Assessment Tool (SWAT). These studies typically make use of rich parcel-level data on land heterogeneity and have examined the economic decisions and environmental implications of tillage (Kurkalova, et al. 2006; Wu, et al. 2004), crop rotation (Wu and Babcock 1998), crop choice (Hendricks, et al. 2014), land allocation (Laukkanen and Nauges 2014), as well as conservation investments for BMP (Rabotyagov, et al. 2010; Rabotyagov, et al. 2014). For example, using sampled points from National Resource Inventory 1982-1997, Wu, et al. (2004) develop a random utility model of crop and tillage choices, and examine the impacts of conservation payments for conservation tillage and crop rotation on agricultural runoffs by linking the economic model with parametric environmental production functions. Rabotyagov, et al. (2014) develop an integrated assessment model, which couples a biophysical SWAT model with optimization algorithm, and identify the most cost-effective subwatersheds to target for cropland conservation investments in order to reduce the extent of hypoxic zone in Gulf of Mexico. While these integrated ecological-economic models provide useful insights regarding the design of agri-environmental policies, they often lack data on farmers' socioeconomic, demographic and behavioral characteristics that are important in farmer decision-making. In addition, to my knowledge, these integrated models do not focus on the heterogeneity in farmers' land management decisions such as fertilizer demand resulting from differentiated land and operator characteristics.

Descriptive Evidence on Heterogeneity in Phosphorus Price Responsiveness

Before I outline the structural model, it is useful to investigate the effects of field-level land and farmer characteristics on phosphorus fertilizer demand, and to examine whether there is descriptive evidence on heterogeneity in the price responsiveness of phosphorus fertilizer demand. Following Gillingham (2014), I run two reduced-form models that aim to quantify the potential heterogeneity in fertilizer demand elasticities, and the results are shown in tables 28 and 29 in the Appendix B. In particular, table 28 presents an ordinary least squares (OLS) regression in which the heterogeneity is introduced by interacting phosphorus fertilizer prices with variables that account for farmers' environmental attitudes and field-level land quality. The results reveal the significance of field-level land and operator characteristics: farmers that are more risk-averse would apply a higher fertilizer rate to avoid potential yield loss, and farmers would apply at a higher rate for fields with good soil quality to maximize the yield potential. While there are no distinctive differences in phosphorus application rates due to crop choices, farmers do over-apply if they follow a multi-year fertilization schedule as opposed to single-year application, which is consistent with observations from agronomists. The results also illustrate the heterogeneity in phosphorus fertilizer demand. The negative and significant coefficient for the variable *Normalized P price * good soil* suggest that farmers have a higher price elasticity when managing a field with high soil quality. This may indicate that good soil quality could serve as a substitute for phosphorus fertilizer, but it may also arise from the fact that the farmers have been applying fertilizers at much higher rates on fields with good soil quality in the first place. Similarly, farmers more familiar with 4R nutrient stewardship have a more inelastic fertilizer demand; however, maybe this is because they are already applying at a lower rate.

Table 29, on the other hand, investigates the heterogeneity in fertilizer demand elasticities by estimating a quantile regression. While linear regression estimates the conditional mean of the fertilizer demand given values of covariates, quantile regressions estimate the conditional median or other quantiles of the dependent variable. Table 29 presents the 25th, median and 75th quantile regression results. The coefficient and implied elasticity on the normalized phosphorus price indicate that at the 25th quantile, of response, the elasticity is -3.003, considerably higher than the estimate for the median quantile regression result of -0.8477 and the 75th quantile regression of -0.5364. The declining magnitude of the price elasticity given the rise of fertilizer prices is intuitive: one farmer cannot always cut back phosphorus applications at high prices because phosphorus fertilizer is an essential nutrient input to the crop production. While the quantile regressions only reveal the heterogeneity in responsiveness, they do not explain the sources of the heterogeneity, unlike the interactions approach shown in table 28.

Finally, while the estimated elasticities at the median and 75th quantile regressions fall within the range of previous estimates, the implied elasticities from the 25th quantile regression and OLS estimation seem a bit high, suggesting possible measurement errors in the reported phosphorus prices in the single-year cross-sectional data, and more careful analysis is needed to identify the elasticity of fertilizer demand. Nonetheless, they provide descriptive evidence on heterogeneity in phosphorus price responsiveness.

Conceptual Framework

In this section I present the microeconomic behavioral model of farmers' profit-maximizing behavior. In this modeling framework, I assume each farmer has only one field[17]. Within the modeling framework and at the field level, a farmer is assumed to make the optimal crop choices as well as the optimal input demand levels that yield the highest expected profit. These choices are made simultaneously, the decisions of input demand and output supply may depend on the crop choices and vice versa. Suppose that for each field, the farmer can choose among C crops and J inputs to maximize the profit. Assume further that farmers consider input and output prices to be exogenous[18] and they could only choose one crop for each field. As a result, the field-level expected profit $\pi_{i|c}$ for field i given crop choice c is given by

$$\pi_{i|c} = [p_{ic}y_{ic} - \sum_j r_{ijc} * x_{ijc} \,|crop = c] \qquad (16),$$

in which p_{ic} is the crop choice farmer i receives, y_{ic} is the expected yield for field i given crop choice c, while r_{ijc} and x_{ijc} denotes the corresponding input price and input quantity for input j. Note that the expected profit $\pi_{i|c}$ is conditional on choice c. The

[17] This assumption is due to the nature of our farmer survey data, in which we ask the farmer to randomly pick one field and answer questions on land use land management for that particular field.

[18] This assumption is reasonable since we focus on farmers in the western Lake Erie basin while the input and output prices for agricultural commodities are commonly driven by national-level or even international-level macroeconomic trends.

unconditional profit function can be written as a weighted average of the conditional profit function:

$$\pi_i = \pi_{i|c} * \varphi_{ic} \qquad (17),$$

where φ_{ic} denotes the probability of farmer i growing crop c in field i and it is a binary variable taking on the value of 0 or 1. Furthermore, I assume each farmer could only choose one crop for one field so that $\sum_{c\in C} \varphi_{ic} = 1$. Equation (17) implies that the profit-maximizing decisions of farmers could be modeled as a two-stage problem: in the first stage, the farmers choose the optimal crop that yields the highest expected utility, and then in the second stage, conditional on the crop choice, the farmers choose the output and input levels to maximize the profit.

Note that this modeling framework only captures the farmers' extensive (crop choice) and intensive margin responses (nutrient input demand) to policies on a single field, and thus ignores the extensive margin changes at the farm level in terms of changes in mix of crop production among different fields in a farm[19]. For example, this model might underestimate the impacts of nutrient management policies if these policies lead to changes in crop mix at the farm level in addition to changes in crop and nutrient applications in the chosen field.

[19] Table 27 in the Appendix B presents results on the effects of changes in input/output prices on mix of crop production at the farm level, in which the crop mix is measured as percentage of corn planted for the farm. I find no statistical evidence that changes in phosphorus fertilizer prices, which could result from alternative nutrient management policies such as fertilizer taxes, would lead to significant changes in crop mix at farm level. This may reflects the fact that changes in crop mix does not occur in the short run and thus may not be captured in our cross-sectional farm survey data. However, this may also suggest my model that focuses on field-level changes in crop choice and input demand and ignores farm-level crop-mix changes could provide a reasonable approximation of the impacts on farm welfare.

Before I discuss the two-stage decision problem in detail, there is one more complication in this model. The phosphorus fertilizer application rates are substantially higher for fields with multi-year application than fields with single-year application, in addition to differences due to crop choices. This is intuitive because farmers may apply a rate that could last for more than a year in multi-year applications. As a result, in the first stage, it is critical to model both the crop choices and the phosphorus application frequency choices. In particular, we model the choices in the first stage as a combination of crop and phosphorus application frequency choices, which include five distinct choices denoted by L: corn and single year application (corn-single, *cs*), corn and multi-year application (corn-multi, *cm*), soybean and single year application (soybean-single, *ss*), soybean and multi-year application (soybean-multi, *sm*) and other crop choices (other, *o*). Let $u_{il}(p_{il}, r_{il}, \boldsymbol{I_i}, \boldsymbol{L_i})$ denotes the utility the farmer gets from choosing the specific crop and fertilizer frequency choice l (l could be $cs, cm, ss, sm\ and\ o$). Because the farmers' preferences are unknown to the researcher, the utility is assumed to be a random variable and written as

$$u_{il}(\boldsymbol{Z_{il}}) = v_{il}(\boldsymbol{Z_{il}}) + \varepsilon_{il} \quad (18),$$

where $v_{il}(\boldsymbol{Z_{il}})$ is the observed mean of $u_{il}(\boldsymbol{Z_{il}})$, ε_{il} is the random error term. $\boldsymbol{Z_{il}}$ represent a set of socioeconomic and land characteristics for the farm, field and operators, as well as output and input prices. $v_{il}(\boldsymbol{Z_{il}})$ is commonly specified as a linear combination of these characteristics, that is $v_{il}(\boldsymbol{Z_{il}}) = \boldsymbol{Z_{il}} * \boldsymbol{\gamma_l}$. Maddala (1983) shows that if the residuals ε_{il} are assumed to be independently and identically distributed with the extreme value distribution, then the probability that the farmer will choose a crop and fertilizer application frequency choice l^* that yields the highest expected utility among the five choices, which can be modeled as a multinomial logit model:

$$P_{il^*} = prob(l = l^*) = \frac{e^{Z_{il^{**}}\gamma_{l^*}}}{\sum_{l \in L} e^{Z_{il^*}\gamma_l}} \qquad l = cs, cm, ss, sm, o \quad (19).$$

Given the first stage crop and phosphorus fertilization frequency choice l^*, the equation (16) now could be rewritten as

$$\pi_{i|l^*} = [p_{il^*} y_{il^*} - \sum_j r_{ijl^*} * x_{ijl^*} \, | l = l^*] \qquad (20)$$

Maximizing the profit function yields the optimal input and output supply functions. I use the dual representation of farm profit maximizing behavior. While a primal approach would require specifying a production function for grains, the dual approach provides a convenient alternative that is based on the specification of a flexible indirect profit function for the second stage, $\pi_{i|l^*}$. The input demand and output supply functions can be derived from the indirect profit function using Hotelling's lemma and will be functions of exogenous output/input prices and quasi-fixed inputs (Chambers 1988).

Estimation Strategy

The Quadratic Profit Function

Any well-behaved profit function must satisfy the following regularity conditions: convexity in prices, homogeneity of degree one in prices, symmetry and monotonicity. I specify a quadratic profit function to flexibly approximate the true profit function. This quadratic profit functional form is chosen over other popular specifications (e.g., generalized Leotief or translog) because it yields directly a set of output supply and input demand equations that make it easier to model farmers' nutrient application decisions and are linear in input and output prices (Arnade and Kelch 2007; Laukkanen and Nauges 2014; Moschini 1988; Shumway and Alexander 1988; Shumway, et al. 1988), and it allows for negative profits [20] (Moschini 1988; Villezca-Becerra and Shumway 1992). The profit and all output and input prices in the profit function are normalized by the labor price, yielding a normalized profit function. I impose the common regularity conditions in this specification and the normalized profit function can be specified as

$$\overline{\pi_{\imath|l^*}} = \beta_l + \alpha_l * \overline{p_{\imath l}} + \sum_j \beta_{jl} * \overline{r_{\imath jl}} + \sum_d \xi_{dl} * z_{idl} + \sum_j \sum_{k\neq j} \sigma_{jkl} * (\overline{r_{\imath jl}} * \overline{r_{\imath kl}}) + \frac{1}{2}\sum_j \gamma_{jl} * \overline{r_{\imath jl}}^2 + \frac{1}{2}\sum_j \varpi_l * \overline{p_{\imath l}}^2 + \sum_j \sum_d \rho_{jdl} * \left(\overline{r_{\imath jl}} * z_{idl}\right) + \sum_j \omega_{jl} * (\overline{r_{\imath jl}} * \overline{p_{\imath l}}) + \sum_d \zeta_{dl} * (z_{idl} * \overline{p_{\imath l}}) + \frac{1}{2}\sum_d \varsigma_{dl} * z_{idl}^2 \quad (21),$$

where $\overline{p_{\imath l}}$ and $\overline{r_{\imath jl}}$ denotes the prices for output l such as corn or soybean and variable input j such as phosphorus, nitrogen fertilizer prices the farmer received, and z_{iml} denote the quasi-fixed input d which in my case are land and machinery. There are two things worth noting in equation (21): first, the upper bar in the equation denotes the normalized profit $\overline{\pi_{\imath|l^*}}$, output price $\overline{p_{\imath l}}$ and input price $\overline{r_{\imath jl}}$ for variable input j. Second, I assume the farmers take input and output prices as exogenous and thus the normalized profit function is a function of exogenous prices for output and variable inputs – phosphorus ($j = P$), nitrogen ($j = N$), or manure ($j = M$), and quantities for quasi-fixed inputs and their interaction terms. $\beta_l, \alpha_l, \beta_{jl}, \xi_{ml}, \sigma_{jkl}, \gamma_{jl}, \varpi_l, \rho_{jdl}, \omega_{jl}, \zeta_{dl}, \varsigma_{dl}$ are the parameters needs to be estimated.

[20] The translog specification requires logarithmic transformation of profit and thus does not allow restricted profit to be negative.

Following Hotelling's lemma, the output supply function for crop l can be obtained by differentiating the normalized profit function with the output price $\overline{p_{\iota l}}$. As a result, differentiating both equations (20) and (21) by $\overline{p_{\iota l}}$ yields:

$$y_{il} = \alpha_l + \varpi_l * \overline{p_{\iota l}} + \sum_j \omega_{jl} * \overline{r_{\iota j l}} + \sum_d \zeta_{dl} * z_{idl} \quad (22).$$

Similar, the variable input demand equation for input j ($j = P, N, M$) can be obtained by differentiating the profit function with respect to the input price $\overline{r_{\iota j l}}$:

$$-x_{ijl} = \beta_{jl} + \gamma_{jl} * \overline{r_{\iota j l}} + \sum_{k \neq j} \sigma_{jkl} * \overline{r_{\iota k l}} + \omega_{jl} * \overline{p_{\iota l}} + \sum_d \rho_{jdl} * z_{idl} \quad (23),$$

The derivation so far follows the standard quadratic profit function and the resulting output supply and input demand equations (22)-(23) are functions of only exogenous prices and quasi-fixed inputs. However, previous economic and agronomic studies have shown that crop and nutrient management choices depend critically on heterogeneous field-level operator and land characteristics. As a result, I make some modifications in the profit function to allow for the heterogeneous effects due to the field-specific land and operator characteristics. Mathematically, this is equivalent to explicitly expressing the intercepts in the output supply α_l and the input demand β_{jl} as functions of these heterogeneous attributes:

$$\alpha_l = \alpha_{l0} + \alpha_{l1} * \boldsymbol{I_i} + \alpha_{l2} * \boldsymbol{L_i} \quad (24)$$

$$\beta_{jl} = \beta_{jl0} + \beta_{jl1} * \boldsymbol{I_i} + \beta_{jl2} * \boldsymbol{L_i} \quad (25)$$

In equations (24) and (25), $\boldsymbol{I_i}$ and $\boldsymbol{L_i}$ represents the individual-level operator and field characteristics respectively, and α_{l0} and β_{jl0} are the new intercepts.

In addition and as I discussed earlier in the introduction, farmers may respond to fertilizer price changes differently in fields whose land characteristics may serve as substitutes for fertilizer than otherwise, while farmers with different levels of information sets such as nutrient stewardship are expected to behave differently as well. As a result, I use four variables, which represent land quality and familiarity with nutrient stewardship, to account for the heterogeneity in price elasticity of fertilizer demand. In particular, these four variables are denoted as $\boldsymbol{S_i}$, which include three soil quality variables (dummy for top soil, dummy for poor soil, dummy for steep slope) and one variable that account for the familiarity with nutrient stewardship (familiar_4R). For simplicity, I only introduce these four variables in the phosphorus demand equation, so the coefficient for the phosphorus fertilizer prices can be rewritten as

$$\gamma_{Pl} = \gamma_{Pl0} + \sum_n \eta_{Pln} * S_{in} \qquad (26).$$

In sum, we can obtain the phosphorus, nitrogen fertilizer and manure input demand equation in the following by plugging in equations (25) (26) into equation (23).

$$-x_{iPl} = (\beta_{Pl0} + \beta_{Pl1} * \boldsymbol{I_i} + \beta_{Pl2} * \boldsymbol{L_i}) + (\gamma_{Pl0} + \sum_n \eta_{Pln} * S_{in}) * \overline{r_{iPl}} + \sum_{k \neq P} \sigma_{Pkl} * \overline{r_{ikl}} + \omega_{Pl} * \overline{p_{il}} + \sum_d \rho_{Pdl} * z_{idl} \quad (27),$$

$$-x_{iNl} = (\beta_{Pl0} + \beta_{Pl1} * \boldsymbol{I_i} + \beta_{iPl2} * \boldsymbol{L_i}) + \gamma_{iNl} * \overline{r_{iNl}} + \sum_{k \neq N} \sigma_{iNkl} * \overline{r_{ikl}} + \omega_{Nl} * \overline{p_{il}} + \sum_d \rho_{iNdl} * z_{idl} \quad (28),$$

$$-x_{iMl} = (\beta_{Pl0} + \beta_{Pl1} * \boldsymbol{I_i} + \beta_{Pl2} * \boldsymbol{L_i}) + \gamma_{Ml} * \overline{r_{iMl}} + \sum_{k \neq M} \sigma_{Mkl} * \overline{r_{ikl}} + \omega_{Ml} * \overline{p_{il}} + \sum_d \rho_{Mdl} * z_{idl} \quad (29).$$

Similarly, I could re-write the output supply equation (22) as

$$y_{il} = (\alpha_{l0} + \alpha_{l1} * \boldsymbol{I_i} + \alpha_{l2} * \boldsymbol{L_i}) + \varpi_l * \overline{p_{il}} + \sum_j \omega_{jl} * \overline{r_{ijl}} + \sum_d \zeta_{dl} * z_{idl} \qquad (30).$$

The four above equations (27)-(30) thus form a system of input demand and output supply equations[21] that need to be estimated, and these parameters further restricts the parameters in the normalized profit function.

Reduced-form Panel Regression

To quantify the impacts of nutrient management policies on farmer welfare through their effects on phosphorus application, it is critical to accurately estimate how farmers respond to changes in fertilizer price changes induced by alternative policies. In other words, it is important to identify the price elasticity of phosphorus fertilizer demand. Timmins and Schlenker (2009) and Chetty (2008) argues that reduced-form studies can be used to identify key parameters of interest and then use it in structural models to simulate policy responses by economic agents. Although I did not explicitly represents changes in welfare as a function of the "sufficient statistic" as in Chetty (2008), I estimate a reduced-form panel data model using phosphorus application rates under actual and two hypothetical price scenarios, and identify the mean elasticity of phosphorus fertilizer demand, which is then constrained in the estimation of the structural model.

The data I use in this study comes from a farmer survey in the western Lake Erie basin soliciting farmers' crop and nutrient management choices in 2013, so it is based on data from a single year and a relatively small region. There is some variation in the fertilizer prices paid among the farmers; however, it may not provide enough variation to reveal

[21] The normalized profit function is omitted from the estimation system because first, the parameters in the normalized profit function shown in equation (21) are already estimated in these four input demand and output supply equations, and second, the full covariance matrix of the estimation system containing the restricted normalized profit function tends to be singular due to collinearity (Shumway and Alexander 1988).

the farmers' true demand elasticity of phosphorous fertilizers: over the past decade, the average U.S. phosphorus price index ranges from $300/ton to $900/ton.

As a result, in addition to one question about the farmers' actual fertilizer application rate and fertilizer price paid, I ask two hypothetical questions to induce farmers' responses under alternative phosphorus fertilizer price scenarios. Specifically I ask "if commercial phosphorus fertilizer prices had been \$$X$/ton, what rate of P would you have applied on this field for this most recent crop? ______ lbs/acre", in which X could be 200, 250, 300, 350, 450, 500, 550, 750, 800, 850, and 900. With these two hypothetical questions on P rate and prices in addition to the question on the observed levels, I now have a short panel of three choices and thus could formulate a panel data fixed effects model:

$$-x_{iPlt} = \kappa_{Pl0} + \gamma_{Pl0} * \overline{r_{iPlt}} + \theta_{il} \qquad t = 1,2,3 \quad (31),$$

where θ_{il} is individual fixed effects.

Following equation (31) and for each crop and fertilization frequency choice l, I could estimate the key parameter of interest $\widehat{\gamma_{Pl0}}$ – the mean coefficient for phosphorus fertilizer prices without heterogeneity[22], and this estimated parameter implies a mean elasticity of phosphorus fertilizer demand.

The reduced-form panel data regression shown in equation (31) is similar in spirit, in which the key parameter of interest – mean coefficient for phosphorus fertilizer prices

[22] I account for the uncertainty in this estimate by conducting a series of robustness checks, including a quadratic specification as opposed to a linear phosphorus price, constraining the coefficient to be $\widehat{\gamma_{Pl0}}$ plus or minus the estimated standard error in the structural model estimation. The implied mean elasticity are similar to the main specification. Currently I am also working on examining the stableness of the coefficients in the structural model by running boostrapping regressions with the constrained coefficient randomly drawn from the distribution of $\widehat{\gamma_{Pl0}}$.

$\widehat{\gamma_{Pl0}}$ – is estimated using reduced form regressions and then become constrained in the structural estimation of the four-equation system of output supply and input demand.

Selectivity and Iterative SUR

As shown previously in equation (17), the second stage nutrient management decisions are heavily influenced by the first-stage crop and fertilization frequency choices. For example, the phosphorus fertilizer rates are much higher for corn fields with multi-year applications compared to other fields. As a result, a direct estimation of second-stage input demand and output supply equations while ignoring the first-stage choices would be biased because of sample selection bias. Lee (1983) suggests a two-stage method for estimating equation (18) to correct the sample section bias: first, the multinomial logit model like equation (19) is estimated, and then the estimated probability $\widehat{P_{\iota l}}$ from equation (20) could be used to calculate a Heckman-style inverse Mills ratio:

$$\widehat{\lambda_{\iota l}} = \frac{\phi[\Phi^{-1}(\widehat{P_{\iota l}})]}{\widehat{P_{\iota l}}} \qquad l = cs, cm, ss, sm, o \quad (32).$$

In the second stage, these inverse Mills ratios are added to the equations (27)-(30). For the two-stage selectivity model to work, we need variables that satisfy the exclusion restriction, that is, variables that only enter the first stage crop and fertilization frequency choices but not the second stage field-level nutrient management decisions. Specifically, five variables are used: farm acres, precipitation, percentage of corn grown at the farm level[23], and two dummies for previous year crop choice. The *percentage of corn grown at*

[23] More exogenous variable such as the percentage of high quality soil at the farm level is also used alternatively as robustness check.

the farm level represents the extensive-margin crop-mix at the farm level and controls for the impacts of farm size on operators' ability to apply fertilizers or choose a crop that requires intensive care in a timely fashion; and *previous crop dummies* represents the effects of crop rotation. Combining equation (32) that accounts for the selectivity and equation (31) that provides a reduced-form estimate for $\widehat{\gamma_{Pl0}}$, we can re-write the estimating system of four input demand and output supply equations as follows:

$$-x_{iPl} = (\beta_{Pl0} + \beta_{Pl1} * \boldsymbol{I_i} + \beta_{Pl2} * \boldsymbol{L_i}) + (\widehat{\gamma_{Pl0}} + \sum_n \eta_{Pln} * S_{in}) * \overline{r_{iPl}} + \sum_{k \neq P} \sigma_{Pkl} * \overline{r_{ikl}} + \omega_{Pl} * \overline{p_{il}} + \sum_d \rho_{Pdl} * z_{idl} + \mu_{Pl} * \widehat{\lambda_{il}} + \epsilon_{iPl} \quad \text{(33a)},$$

$$-x_{iNl} = (\beta_{Pl0} + \beta_{Pl1} * \boldsymbol{I_i} + \beta_{iPl2} * \boldsymbol{L_i}) + \gamma_{iNl} * \overline{r_{iNl}} + \sum_{k \neq N} \sigma_{iNkl} * \overline{r_{ikl}} + \omega_{Nl} * \overline{p_{il}} + \sum_d \rho_{iNdl} * z_{idl} + \mu_{Nl} * \widehat{\lambda_{il}} + \epsilon_{iNl} \quad \text{(33b)},$$

$$-x_{iMl} = (\beta_{Pl0} + \beta_{Pl1} * \boldsymbol{I_i} + \beta_{Pl2} * \boldsymbol{L_i}) + \gamma_{Ml} * \overline{r_{iMl}} + \sum_{k \neq M} \sigma_{Mkl} * \overline{r_{ikl}} + \omega_{Ml} * \overline{p_{il}} + \sum_d \rho_{Mdl} * z_{idl} + \mu_{Ml} * \widehat{\lambda_{il}} + \epsilon_{iMl} \quad \text{(33c)}.$$

$$y_{il} = (\alpha_{l0} + \alpha_{l1} * \boldsymbol{I_i} + \alpha_{l2} * \boldsymbol{L_i}) + \varpi_l * \overline{p_{il}} + \sum_j \omega_{jl} * \overline{r_{ijl}} + \sum_d \zeta_{dl} * z_{idl} + \mu_{yl} * \widehat{\lambda_{il}} + \epsilon_{iyl} \quad \text{(33d)}.$$

The above four equations form the final estimating system, and the terms ϵ_{iPl} to ϵ_{iyl} are idiosyncratic error terms, which are possibly correlated across equations and by assumption of mean zero. There are three estimation issues that need to be addressed. First, the joint nature in nutrient management decisions with regard to multiple variable inputs, as evidenced by the correlation among ϵ_{iPl} and ϵ_{iyl}, need to be accounted for. As a result, I use iterative Seemingly Uncorrelated Regression (SUR) to estimate the equations (33a) – (33d) as a whole. Second, the sample selection bias can be tested by investigating whether μ_{Pl}, μ_{Nl}, μ_{Ml}, and μ_{yl} equals zero. If they are statistically significantly different from zero, then the estimates without accounting for the sample selection bias would be biased. Finally, the two-stage method proposed by Lee (1983) only provide unbiased estimates but not the variances (Wu and Babcock 1998). In other words, the estimated standard errors from iterative SUR are biased and thus I report the consistent standard errors estimated from bootstrapping method instead.

Data

	Corn-single	Corn-multi	Soybean-single	Soybean-multi	Other crop
P application rate	64.969	81.038	67.410	83.093	64.292
	(74.747)	(101.348)	(81.864)	(87.515)	(98.408)
N application rate	148.745	149.321	28.039	51.732	93.583
	(109.218)	(114.306)	(66.475)	(109.316)	(125.654)
P fertilizer price	576.931	576.033	574.646	583.803	578.563
	(104.120)	(112.498)	(98.387)	(101.703)	(105.481)
N fertilizer price	235.054	287.384	44.989	77.910	152.990
	(310.014)	(332.681)	(147.549)	(208.695)	(260.525)
# obs	708	368	248	135	96

Table 15. Fertilizer Application Rates and Fertilizer Prices Across Different Alternatives
Note: standard deviations in the parenthesis[24].

Figure 7 shows our study region – the western Lake Erie basin and the Maumee River watershed in particular, which is the largest tributary in Lake Erie and the largest drainage basin in the Great Lakes region. More importantly, the Maumee River

[24] The group t-tests reveal that there is no statistical difference in phosphorus application rates for corn, soybean or other crops when they are applying phosphorus in a single-year fashion. However, the phosphorus application rates are statistically lower for corn or soybean fields with single-year application than fields that apply phosphorus fertilizers in a multi-year frequency.

watershed has been implicated as the largest source of phosphorus flows into Lake Erie: the dissolved reactive phosphorus loadings have increased by over 200 percent from 1995 to 2011, which has been a major cause of harmful algal blooms and other water quality problems in Lake Erie (Michalak, et al. 2013). As part of a NSF-funded project, my collaborators and I conducted a mail survey of 7,500 farmers in the Maumee River watershed on their field, farm and operator characteristics in February - April 2014. We also solicited field-specific responses on crop choices, fertilizer application, and other nutrient management practices in 2013. The addresses of all farmers in the Maumee River watershed were provided by a private vendor, and were pulled from lists of farmers receiving government payments and from farming magazine rolls. The two-round survey was conducted following Dillman's Tailored Design method (Dillman 2011). The total set of mailings included an announcement letter, a survey packet, a reminder letter and a replacement packet for non-responders. The respondents receive a $1 bill in the mailings as an incentive to increase the response rate. Several months before the initial mailings of the survey it was pilot tested using farmers recruited by local extension professionals.

A total of 3,234 surveys were initially returned, of these 438 were no longer farming and 32 surveys did not answer sufficient number of questions. In total, we obtained 2,764 valid survey responses, yielding a response rate of 36.9%. Of these, 1,213 respondents did not provide answers to either fertilizer rate or price questions or certain field or operator characteristics. As a result, a total of 1,551 surveys are used in this analysis and the majority of them are from corn or soybean growers.

Maumee River Watershed

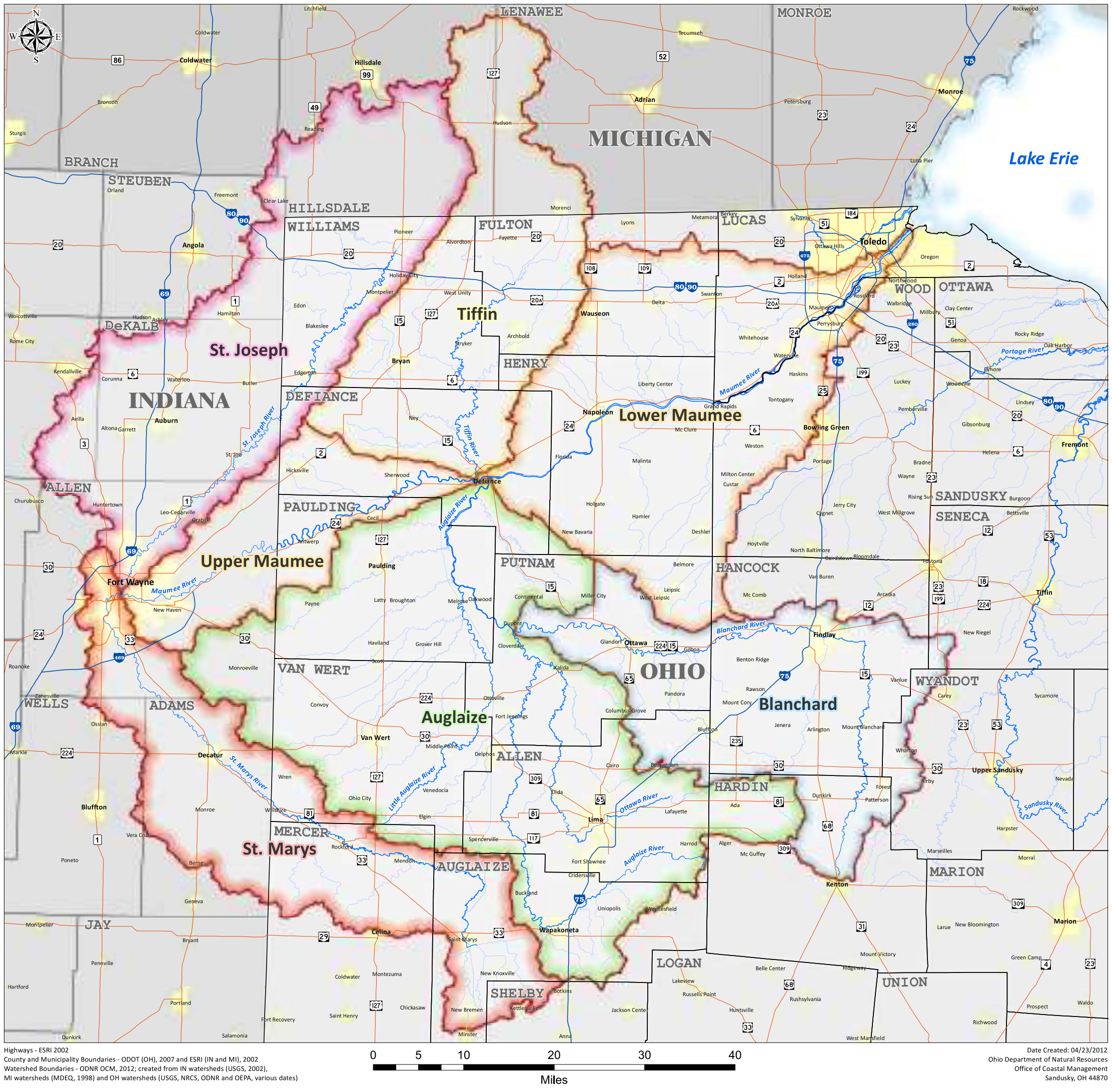


Figure 7. The Maumee River Watershed in the Western Lake Erie Basin

Table 15 shows the phosphorus and nitrogen fertilizer application rates and the prices farmers paid in 2013 by different crop and fertilization frequency choices. Corn requires much higher rates of nitrogen fertilizer than soybean, while fields with multi-year

Variable	Obs	Mean	Std. Dev.	Min	Max
Output Supply and Input Demand					
Expected yield in 2013	1551	125.2449	63.89199	0	250
Phosphorus application rate	1551	70.65575	85.76222	0	1250
Nitrogen application rate	1549	117.6531	116.358	0	1250
Manure quantity	1545	703.3537	6168.947	0	200000
Crop in 2012 is corn	1551	0.315925	0.465033	0	1
Crop in 2012 is soybean	1551	0.524178	0.499576	0	1
Output and Input Prices					
Corn price in 2012	1410	511.0346	104.1846	0	553.2105
Soybean price in 2012	1410	1184.047	152.6893	0	1229.5
Normalized P fertilizer price	1551	57.70151	10.52109	0.02	120
Normalized N fertilizer price	1547	38.52827	58.67502	0	433.3333
Manure price	1545	0.001053	0.006346	0	0.08
P price norm * familiar_4R	1542	92.59627	75.77997	0	360
P price norm * slope	1551	21.35525	28.41432	0	90
P price norm * top soil	1551	20.7746	28.49471	0	120
P price norm * poor soil	1551	18.15941	27.38308	0	110
Farmer Characteristics					
Familiar with 4R	1542	1.597925	1.254897	0	4
Risk attitudes (10 = risk loving)	1537	5.592713	2.347849	0	10
Age	1520	58.51908	17.29406	5	548
Education	1527	3.027505	1.311019	1	6
Years of farming experience	1525	38.3082	17.17624	3	400
Female operator	1527	0.008513	0.091905	0	1
Farm income	1414	3.204385	1.305965	1	5
Field Characteristics					
Field acres	1551	50.864	65.46771	0	1100
The field is rented	1538	0.363459	0.481152	0	1
Distance (meter) to Lake Erie	1366	96855.32	43155.65	3685	191613
Soil texture is clay	1551	0.221148	0.415154	0	1
Soil texture is sand	1551	0.021277	0.144351	0	1
Field is Highly Erodible Land	1534	0.249022	0.577279	0	2
Field has a slope > 2%	1551	0.372663	0.483669	0	1
Field has good soil	1551	0.359768	0.480087	0	1
Field has poor soil	1551	0.317215	0.465542	0	1

Continued

Table 16. Summary Statistics of Field, Farm, and Farmer Characteristics

Table 16 continued

Farm Characteristics					
Farm acres	1548	573.0949	771.5575	0	7050
% corn in all planted acres	1376	0.406479	0.490897	0	10.53333
Farm has livestock	1505	0.304319	0.460271	0	1
Enrolled in crop insurance	1530	0.701961	0.457546	0	1

application frequency have a much higher phosphorus fertilizer rates than fields that apply fertilizers each year. The differences in the nitrogen prices resulting from different forms of nitrogen fertilizer used for corn and soybean: urea was more commonly used for soybean while corn used more ammonia and they have different nitrogen content[25]. Table 16 shows the summary statistics for the farmer survey, including crop and nutrient application rates, output and input prices, field characteristics, farm characteristics, and operator characteristics. A comparison between our data and the Census of Agriculture data for counties in the Maumee River watershed reveals that our sample is skewed toward large farms with high gross sales and farmers who additionally earn off-farm income. Most of the variables in table 16 are intuitive; I just want to highlight one group of variables – the interaction terms between the normalized phosphorus fertilizer prices and the four variables that control for heterogeneous responses due to different soil quality and familiarity with 4R Nutrient Stewardship[26].

[25] The nitrogen fertilizer prices are adjusted for the percentage of nitrogen content.
[26] 4R refers to using the Right Source of nutrients at the Right Rate and Right Time in the Right Place.

Results and Discussion

Table 17 shows the results for the first-stage crop and fertilization frequency choices. Relative to other crop choices, fields with a larger size and better soil quality have a higher probability of choosing corn or soybean. Farmers currently enrolled in crop insurance program[27] or with higher farm income are more likely to grow corn, while farmers who rent a field have a higher probability to choose soybean. Many other characteristics do not have statistically significant effects, suggesting that farmers in our study region may follow a historic crop rotation pattern as evidenced by the significance for previous crop dummy. I still model crop choice in the first stage because according to agronomists phosphorus applications depend more on crop choice on a particular year[28] and the effect of crop rotation is at least in part accounted for by modeling the fertilization frequency in the first stage as well.

Table 18 reports the results for the reduced-form panel data analysis shown in equation (31). This model is estimated separately for each crop and fertilization frequency choices. The mean phosphorus application rates are copied from table 15, and the coefficient for *p_price _norm* is the estimated $\widehat{\gamma_{Pl0}}$ in equation (31) and become the constrained coefficient in the structural estimation for the phosphorus input demand equation. Then the mean estimated elasticity is derived from this estimate while holding all other variables constant at means. On average, the derived elasticity of phosphorus fertilizer

[27] The contemporary crop insurance participation might be endogenous, and thus I ran two robustness checks using county-level yield protection insurance rates in 2012 or historical farmers' loss ratios as instruments and the results are qualitatively similar.
[28] For example, phosphorus application rates for soybean fields and corn fields both in corn-soybean rotation could have different phosphorus application rates.

	corn single	corn multi	soybean single	soybean multi
Previous Crop Choices				
2012 crop is corn	2.2336**	1.7330**	4.0188***	4.0946***
	(0.7992)	(0.8173)	(0.8241)	(0.8624)
2012 crop is soybean	0.0079	-0.0990	-0.3574	-0.6500
	(0.3576)	(0.3777)	(0.4245)	(0.5250)
Input and Output Prices				
P fertilizer price	-4.05E-05	0.0001	-0.0011	0.0022
	(0.0012)	(0.0013)	(0.0014)	(0.0018)
N fertilizer price	0.0011*	0.0015***	-0.0026***	-0.0009
	(0.0005)	(0.0005)	(0.0007)	(0.0007)
Corn price 2012	-0.0015	0.0002	-0.0025	-0.0010
	(0.0020)	(0.0021)	(0.0021)	(0.0024)
Soybean price 2012	0.0016	0.0007	0.0030**	0.0011
	(0.0011)	(0.0012)	(0.0015)	(0.0014)
Field Characteristics				
Poor soil	-0.9097**	-1.1490	-0.6979*	-0.8304*
	(0.3629)	(0.3815)	(0.4029)	(0.4446)
Top soil	0.5194	0.4437	-0.1799	0.1756
	(0.4229)	(0.4352)	(0.4694)	(0.5036)
Precipitation	0.0012	-0.0806	0.0006	0.0428
	(0.0770)	(0.0892)	(0.0977)	(0.0892)
Field acres	0.0208***	0.0200***	0.0167**	0.0182**
	(0.0070)	(0.0070)	(0.0074)	(0.0074)
Distance to Lake Erie	2.56E-06	6.58E-06	3.95E-06	2.84E-06
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Slope	0.0863	-0.0131	0.0061	-0.0306
	(0.1173)	(0.1255)	(0.1300)	(0.1459)
Soil texture is clay	-0.5975*	-0.1392	0.2179	0.5949
	(0.3402)	(0.3528)	(0.3764)	(0.4126)
Soil texture is sand	0.2450	1.0549	0.3793	-0.5422
	(1.1616)	(1.1666)	(1.3038)	(1.5717)
Field is rented	0.1359	0.3494	0.8272**	0.8047**
	(0.3401)	(0.3527)	(0.3784)	(0.4105)

Continued

Table 17. First Stage Multinomial Logit Model of Crop and Fertilizer Application Frequency Choices

Table 17 continued

Farmer Characteristics				
Age	-0.0003	0.0057	-0.0012	-0.0045
	(0.0137)	(0.0154)	(0.0157)	(0.0175)
Familiar with 4R Nutrient	-0.1302	-0.1719	-0.2491*	-0.3317**
Stewardship	(0.1251)	(0.1301)	(0.1419)	(0.1545)
More risk loving	-0.0858	-0.0046	-0.0070	-0.1384*
	(0.0645)	(0.0679)	(0.0730)	(0.0793)
Education	-0.1017	0.0996	0.0107	-0.0316
	(0.1076)	(0.1115)	(0.1212)	(0.1360)
Years of farming experience	-0.0003	-0.0050	-0.0075	-0.0065
	(0.0069)	(0.0094)	(0.0082)	(0.0093)
Female operator	10.618	11.111	10.87	12.147
	(412.83)	(412.83)	(412.83)	(412.83)
Farm income	0.3632**	0.3884**	0.2139	0.3322*
	(0.1609)	(0.1676)	(0.1809)	(0.1926)
Farm Characteristics				
Has crop insurance	0.7918**	0.8517**	0.5870	0.4722
	(0.3154)	(0.3354)	(0.3576)	(0.4034)
Farm acres	-0.0002	-0.0001	-0.0002	3.38E-06
	(0.0002)	(0.0002)	(0.0003)	(0.0003)
% corn in farm acres	0.1374	0.1579	-3.481***	-0.4756
	(0.3073)	(0.3162)	(0.7182)	(0.4962)
Intercept	-0.8430	-0.7509	-0.9311	-3.6794
	(2.5433)	(2.8446)	(3.2373)	(3.1296)
# Observations	707	368	248	135
Log-likelihood	-1194.24			
Pseudo R2	0.1917			

Note: *, **, and *** indicates the coefficient is significant at 10%, 5% and 1% level, respectively. The four choices are classified based on a combination of crop choices in 2013 and phosphorus fertilizer application frequencies: for example, corn-single represents that corn was grown in 2013 and phosphorus was applied on single year basis. The baseline reference group is other all crop choices including wheat and hay.

demand ranges from -0.264 to -0.488. For example, there is a 2.64% reduction in phosphorus fertilizer rate given a 10% fertilizer price increase for corn fields with single-year fertilization. These estimates are similar to previous estimates of elasticity of fertilizer demand (Griliches 1959; Pitt 1983), which ranges from -0.20 to -0.95. A comparison of the elasticity across different fertilization frequency choices reveals that fields with multi-year fertilization application have a significantly higher elasticity of phosphorus demand than fields with single-year application. This makes sense because farmers are more likely to over-apply nutrients under multi-year applications and could make flexible changes facing input price shocks. With this elasticity and mean application rate, I also report number of pounds of phosphorus reduction for a 50% phosphorus price increase, which ranges from 8.8 lbs/acre to 21.25 lbs/acre.

The quadratic normalized profit function leads to a system of output supply and input demand equations that are linear in input prices, which results in the linear specification in the reduced-form panel data model. To account for the uncertainty resulting from the linear specification assumption, I present two robustness checks for the reduced-form panel data model in panels (II) and (III) of table 18. In particular, panel (II) only uses responses from these two hypothetical fertilizer application rates questions and assess the effects of potential "hypothetical bias" on the estimated coefficient in phosphorus fertilizer prices. The implied elasticities are very similar with the main specification except for the corn with multi-year applications, which is also within the range of previous estimates from the literature. In addition, I estimate a quadratic fertilizer demand model and present it in the panel (III) of table 18. Although the estimated regression

	corn single	corn multi	soybean single	soybean multi
Linear panel fixed effects model				
Actual and hypothetical	-0.4376*	-0.5634***	-0.4104***	-0.8462***
phosphorus price	(0.2259)	(0.1689)	(0.1111)	(0.2325)
Intercept	115.89***	112.47***	109.52***	148.71***
	(12.77)	(9.43)	(6.186)	(13.39)
Number of observations	1752	1097	603	405
Implied mean elasticity	-0.2714*	-0.388***	-0.2638***	-0.4876***
Implied average P reduction (lbs/ac) for a 10% price increase	-2.474*	-3.146***	-2.286***	-4.874*
Linear panel fixed effects model – Hypothetical questions only				
Hypothetical	-0.4682***	-0.3616***	-0.3561***	-0.8307***
phosphorus price	(0.1554)	(0.1063)	(0.1012)	(0.2620)
Intercept	124.65***	100.82***	112.63***	155.93
	(8.71)	(5.84)	(5.559)	(14.990)
Number of observations	1168	731	402	270
Implied mean elasticity	-0.2665***	-0.2456***	-0.2101***	-0.4383***
Implied average P reduction (lbs/ac) for a 10% price increase	-2.623***	-1.988***	-1.956***	-4.752***
Quadratic panel fixed effects model				
Actual and hypothetical	-2.9534*	-1.5802	-2.0950**	-2.3991**
phosphorus price	(1.720)	(1.204)	(0.8268)	(0.9850)
Phosphorus price squared	0.0230	0.0093	0.0155**	0.0140*
	(0.0142)	(0.0097)	(0.0074)	(0.0077)
Intercept	174.2***	136.04***	148.57***	185.91***
	(45.68)	(32.79)	(20.15)	(29.36)
Number of observations	1752	1097	603	405
Implied mean elasticity	-0.249*	-0.395***	-0.256***	-0.4809***
Implied average P reduction (lbs/ac) for a 10% price increase	-2.010*	-3.03***	-2.051***	-4.529*
Average actual phosphorus application rate (lbs/ac)	64.97	81.04	67.41	87.09

Table 18. Estimated Elasticity of Phosphorus Fertilizer Demand from Reduced-form Panel Data Estimation

coefficients are different, the implied mean elasticity of fertilizer demand is almost identical with the main specification. These two robustness checks lend support to the stability of my results shown in the main specification in panel (I).

Table 19 presents the results for the structural model of output supply and input demand equations. Only results on the phosphorus fertilizer demand are shown in table 19, while the results on yield, nitrogen demand and manure demand are shown in the Appendix B Table 30. These models are estimated following equations (33) and separately for each first-stage crop and fertilization frequency choices. First note that the inverse Mills ratio is statistically significant for corn and multi-year application as well as soybean with single year application, suggesting that it is critical to control for selectivity of crop and fertilization frequency choices at least for these two choices. Second, the coefficient for normalized phosphorus fertilizer price is constrained from the estimated $\widehat{\gamma_{Pl0}}$ shown in table 18 panel (I).

The variables of interest are the four interaction terms between fertilizer prices and soil quality and environmental attitudes variables. A comparison between corn and soybean shows that corn growers would apply phosphorus fertilizers differently according to land and behavioral characteristics, while soybean growers tend to apply fertilizers in a more similar manner in all fields. This is because that soybean typically requires less fertilizers compared to corn so less costs are incurred, and the yield gains for soybeans due to phosphorus fertilizer applications are smaller compared to corn. In addition, heterogeneity in fertilizer demand is more evident for multi-year applications.

	corn single	corn multi	soybean single	soybean multi
Constrained P Fertilizer Prices Estimated from Reduced-form Panel Data Model				
Actual and hypothetical	-0.4376*	-0.5634***	-0.4104***	-0.8462***
phosphorus price	(0.2259)	(0.1689)	(0.1111)	(0.2325)
Phosphorus Fertilizer Prices with Interactions - Targeting				
Normalized P price * familiar 4R	0.4911**	0.7447**	-0.1678	-0.0778
	(0.2588)	(0.3279)	(0.5537)	(1.1047)
Normalized P price * slope	-1.0938*	-2.2810	0.8596	-7.9736
	(0.6278)	(1.4685)	(1.6485)	(7.0200)
Normalized P price * top soil	0.2375	-3.8295***	-1.6521	5.5019**
	(0.6959)	(0.8599)	(2.0007)	(2.5978)
Normalized P price * poor soil	-1.0487	-0.1748	-1.1463	1.6373
	(0.6595)	(1.5634)	(0.8441)	(1.8262)
Input and Output Prices				
Normalized N price	0.1492**	0.1344	0.1259	-0.0626
	(0.0587)	(0.1090)	(0.2169)	(0.2094)
Normalized expected corn price	0.3813	-1.0552*	0.2257	-2.123***
	(0.3054)	(0.5675)	(0.5836)	(0.7496)
Normalized manure price	-3207.8	8100.99	-34990	24033
	(5403.8)	(6454.0)	(63886)	(27804)
Farmer Characteristics				
Familiar with 4R Nutrient Stewardship	-31.093**	-51.581***	1.2687	-5.245
	(15.444)	(19.746)	(32.340)	(64.905)
More risk loving	-3.648**	-9.737***	-1.4801	5.065
	(1.5678)	(2.8193)	(2.5446)	(3.8733)
Age	-0.268	-1.1536	-0.2003	1.5212
	(0.3591)	(0.8452)	(0.8768)	(1.2030)
Education	-0.9712	-6.2659	-2.9087	-24.328***
	(2.9465)	(5.4643)	(4.6654)	(6.9833)
Years of farming experience	0.2903	0.2568	-0.3569	-2.4311**
	(0.1938)	(0.7366)	(0.7292)	(1.1953)
Female operator	-77.096	-69.208	445.63***	11.745
	(52.748)	(93.813)	(74.779)	(55.686)

Continued

Table 19. SUREG Regression Results for Phosphorus Fertilizer Rate Equation with Bootstrapped Standard Errors

Table 19 continued

	Corn single	Corn-multi	Soybean-single	Soybean-multi
Farm Characteristics				
Farm income	-6.737**	6.1030	3.3666	13.469*
	(3.1446)	(5.7922)	(5.3278)	(7.4660)
Farm has livestock	-14.879**	-13.975	-16.841	13.042
	(7.4146)	(12.755)	(13.089)	(21.358)
Farm has crop insurance	2.152	13.785	-14.890	42.576**
	(8.2328)	(14.359)	(12.528)	(18.849)
Field Characteristics				
Field acres	-0.0218	-0.0997	0.0690	-0.5525**
	(0.0471)	(0.1022)	(0.2174)	(0.2509)
Field is rented	1.799	-4.8941	-5.4209	-1.409
	(7.7259)	(12.783)	(12.849)	(17.944)
Soil texture is clay	21.006**	-9.9542	3.6203	2.3869
	(10.665)	(14.619)	(12.156)	(17.852)
Soil texture is sand	34.987	-23.637	24.238	-156.89**
	(28.740)	(37.490)	(42.519)	(76.702)
Distance (meter) to Lake Erie	-0.0001	-0.0004***	8.23E-05	-3.78E-04*
	(0.0000)	(0.0001)	(0.0001)	(0.0002)
Slope is great than 2%	71.697**	130.62	-58.929	462.936
	(36.515)	(87.470)	(97.175)	(420.30)
Field has good soil	2.7308	225.48***	112.22	-343.31**
	(41.349)	(50.811)	(119.51)	(152.87)
Field has poor soil	72.66*	19.229	74.621	-124.64
	(38.376)	(93.656)	(49.289)	(111.11)
Inverse Mills ratio for 1st stage	-24.147	-52.483*	23.748*	1.5125
crop and P frequency choices	(18.896)	(28.081)	(13.598)	(20.626)
Intercept	141.15***	389.71***	92.206*	314.122***
	(32.600)	(87.104)	(53.310)	(84.130)

Note: *, **, and *** indicates the coefficient is significant at 10%, 5% and 1% level, respectively. Standard errors are included in the parentheses.

Table 30 in the Appendix B shows no statistical evidence of negative yield impacts for an increase in phosphorus fertilizer prices for the range of prices that we analyze. The regression coefficients for phosphorus fertilizer price in the yield equation are negative for corn or soybean fields with multi-year application, however, they are not statistically significant. The nitrogen and manure demand equations reveal some evidence of input substitutability: for example, an increase in phosphorus fertilizer prices would significantly increase the demand for nitrogen, suggesting that phosphorus and nitrogen fertilizers are substitutes, at least for corn fields on a multi-year application schedule. The manure demand equation, on the other hand, shows a counter-intuitive result that manure serve as a supplement for phosphrus application for corn fields with single-year application.

Tables 31 and 32 in the Appendix B show results on two additional robustness checks. In particular, in table 31 I present structural model estimation results directly estimated without constraining the coefficient on phosphorus fertilizer price to be the estimated coefficient from the reduced-form panel data model. Table 32, on the other hand, presents the results that omit the manure demand equation and manure prices due to concerns about the measurement errors in manure quantity demanded and the manure price reported, as evidenced by the abnormally large coefficient for manure prices in table 19. The results omitting manure demand and manure prices yield almost identical price elasticity of phosphorus fertilizer demand, however, directly estimating the model without constraining the coefficients leads to a much higher estimate of phosphorus demand elasticity, except for corn fields with single-year application. For example, the estimated elasticity for corn fields with multi-year application is as high as -2.62, which is aligned with the estimates from the OLS regression and quantile regressions of phosphorus fertilizer demand equations shown in tables 28 and 29 in the Appendix B. As discussed earlier, this estimate is beyond the range of estimated elasticties reported from most previous studies, suggesting that it might be inaccurate and insufficient to quantify the elasticity of fertilizer demand using just cross-sectional data on fertilizer applications in one year. As a result, the results that combine the identification of elasticity using reduced-form panel data model seem more reliable, although more work is needed to reconcile the differences.

To better interpret the heterogeneity in the elasticity, I translate the significant coefficients for the interaction terms in table 19 into semi-elasticities presented in table 20. Specifically, table 20 reveals that the reduction in phosphorus application rate is slightly higher for farmers more familiar with 4R nutrient stewardship, however, the differences may be too small for an education campaign to make a difference in environmental outcome. However, to fully examine the impacts of the education campaign policy, a model is needed to quantify how an education campaign would affect the environmental awareness with 4R nutrient stewardship, and thus results for this policy presented here are descriptive and exploratory in nature. In addition, a comparison between phosphorus reduction on average soil fields and high quality soil fields also suggest that soil quality does serve as substitutes for fertilizer input: farmers would cut more phosphorus on fields with better soil quality when employing multi-year

	Corn-single	Corn-multi	Soybean-single	Soybean-multi
Mean P application rate	64.97	81.04	67.41	87.09
lbs/acre P application reduction for a 50% P price increase				
Mean	-8.8	-15.72	-8.9	-21.25
4R - slightly familiar	-11.15	-16.5		
4R- very familiar	-12.15	-17.95		
Flat land	-12.15			
Steep slope	-11.2			
Average soil		-15.9		-22.5
High quality soil		-16.2		-25.8

Table 20. Heterogeneity in Semi-elasticity of Fertilizer Demand Across Behavioral and Land Characteristics

Note: The lbs/acre reduction in phosphorus application rate for a 50% phosphorus fertilizer price increase is calcualted based on the semi-elascticity implied from the estimated coefficients on Normalized P price and its interactions with four behavioral and land characteristics from Table 19. Blank cell is due to the insignificant regression coefficient on these four interaction terms.

# Scenario	Scenario Description
1	Baseline
2	Education campaign to increase familiarity with 4R
3	P application limit to 100 lbs/acre
4	25% uniform P tax
5	50% uniform P tax
6	50% targeted P tax on steeped sloped land only
7	50% targeted P tax on land in high runoff potential subbasin only
8	50% targeted P tax on land with good soil quality only
9	100% uniform P tax

Table 21. Alternative Nutrient Management Policy Scenarios

fertilization schedules. For example, a 50% increase in phosphorus fertilizer prices would lead to a 25.8 lbs/acre and 22.5 lbs/acre reduction in phosphorus application rates for top soil fields and average soil fields growing soybean, respectively. The differences between the reductions on flat land and steep sloped fields also reveal the substitutability between soil quality and nutrient inputs like fertilizers. These results confirm my conjecture that farmers think that better soil quality could mitigate the potential negative impacts of reduced phosphorus fertilizer.

As discussed earlier, the International Joint Commission has determined that at least 40% reduction in agricultural phosphorus nutrient runoff is needed to restore Lake Erie from a eutrophic state to a lake free from HAB (Reutter, et al. 2011). Table 21 presents eight alternative nutrient management policy scenarios that aim to reduce phosphorus runoff, which include uniform tax, spatially targeted tax based on land characteristics or location of the parcels, education campaign to enhance familiarity with 4R, and a limit on phosphorus application rates. The spatially targeted tax are either only applied to parcels with steep slope or located in the high runoff potential, ecologically sensitive subbasins, which are determined based on subbasin-level phosphorus loading coefficients from a previously calibrated SWAT model (Gebremariam, et al. 2014). By attempting to link the taxes with potential environmental damages at the field level, these spatially-targeted policies are designed to improve the cost-effectiveness of the policies. The familiarity with 4R nutrient stewardship has five levels: not at all (0), slightly (1), moderately (2), very (3) and extremely (4) familiar. I assume the education campaign policy would move the familiarity with 4R up a level for each farmer and induce the farmer to be a little

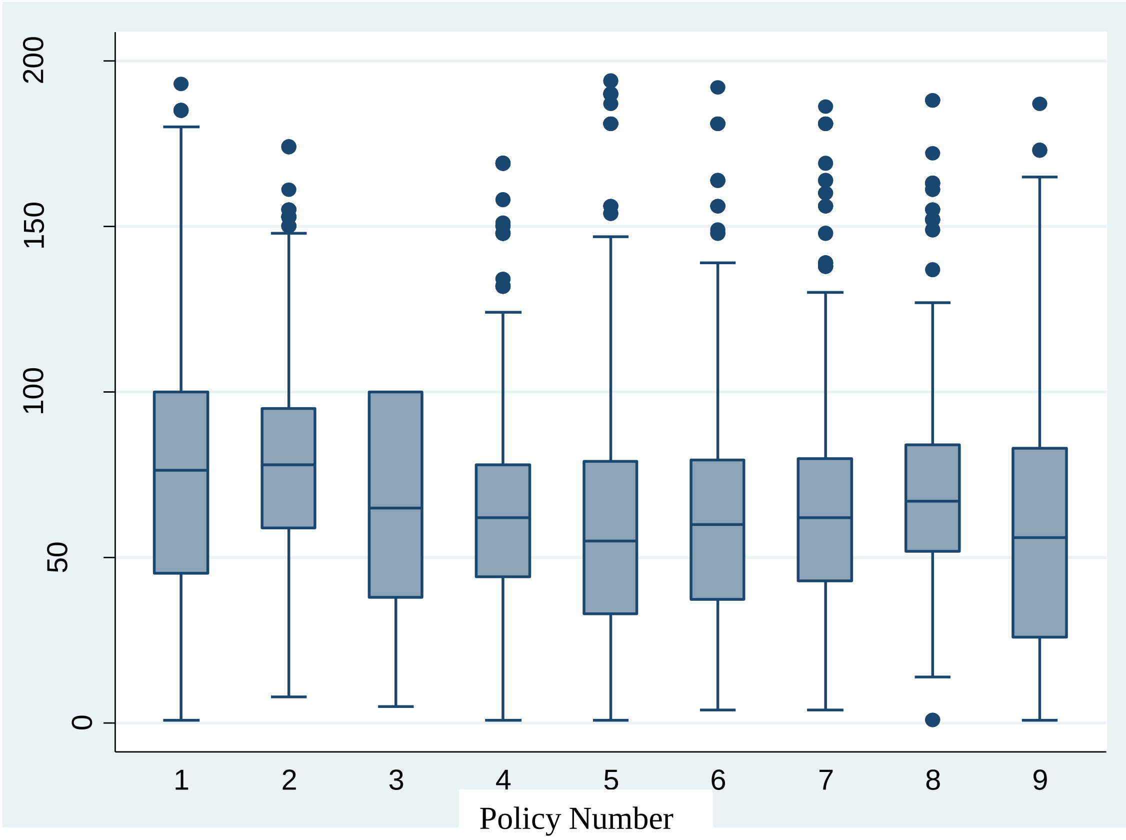


Figure 8. Impacts of Alternative Nutrient Management Policies on Predicted Phosphorus Application Rates at Field Level

Note: The numbers in the x-axis represents different nutrient management policies as shown in Table 21, and they represent: (1) baseline, (2) education campaign, (3) phosphorus limit, (4) 25% uniform P tax, (5) 50% uniform P tax, (6) 50% targeted P tax on steeped slope land only, (7) 50% targeted P tax on fields in high runoff pollution subbasin only, (8) 50% targeted P tax on fields with good soil quality only, and (9) 100% uniform P tax, respectively.

more familiar with 4R. The fertilizer application limit is applied by finding an equivalent farmer-specific fertilizer tax for farmers who apply above the limit following Hansen (2004).

Figure 8 shows the impacts of these eight alternative nutrient management policies on predicted fertilizer application rates at individual field level compared to the baseline scenario. The middle line shows the median application rate, while the upper and lower bar of the blue box denote the 75th and 25th percentile, and the upper and lower bar of the whiskers are two adjacent values[29]. The impacts on phosphorus application reduction from each of the eight alternative policies can be examined by comparing the distribution of phosphorus application rates under that policy and the baseline conditions.

On average, the education campaign does not reduce phosphorus application rate much, and regulatory approaches such as phosphorus fertilizer limit and fertilizer taxes are more effective in reducing fertilizer application rates. A comparison among the three uniform taxes shows that a higher tax rate leads to a greater reduction in phosphorus application rates. However, even a uniform fertilizer tax as high as 100% could only lead to less than 30% reduction in phosphorus application rates. In addition, biophysical models of nutrient flows have also shown that it takes more than one percent reduction in phosphorus fertilizer applications to achieve a one percent reduction in phosphorus runoff into Lake Erie (Michalak, et al. 2013). These suggest that a fertilizer tax alone, even ignoring its political feasibility, could not solve problem of harmful algal blooms in Lake Erie. In other words, it is likely that a mix of policy instruments is necessary to achieve the policy goal of 40% reduction in agricultural phosphorus loadings.

A comparison between uniform fertilizer taxes and the targeted tax based on field or locational characteristics reveals that the uniform taxes could have a slightly greater reduction in phosphorus application rates. This is intuitive because the targeted taxes are only applied to a subset of parcels with steep slope or good soil quality or located in the

[29] Upper and lower bar of the whistler is the most extreme values of within Q75 + 1.5*(Q75 – Q25) and Q75 – 1.5*(Q75 – Q25), where Q75 and Q25 are the 75th and 25th percentile of the distribution.

#	scenario	(a) % P reduction	(b) Average direct cost for farmers ($/acre)	(c) Average direct farmer cost + cost of potential yield loss ($/acre)	(d) Average net policy cost ($/acre)	(e) % P reduction per $ direct cost to farmers	(f) % P reduction per $ direct cost to farmers + yield loss cost	(g) % P reduction per $ net policy cost
1	baseline	0	0	0	0	0	0	0
2	4R education	0.69	1	5.5	1	0.69	0.13	0.69
3	100 lbs/acre P limit	18.55	3.49	45.5	6.51	5.32	0.41	2.85
4	25% uniform P tax	12.15	4.71	29.71	0.47	2.58	0.41	25.80
5	50% uniform P tax	23.92	8.98	50.98	0.90	2.66	0.47	26.64
6	50% target slope P tax	20.41	7.29	37.29	3.28	2.80	0.55	6.22
7	50% target subbasin P tax	18.61	2.18	32.18	0.76	8.54	0.58	24.49
8	50% target top soil P tax	9.76	1.75	31.75	0.79	5.58	0.31	12.35
9	100% uniform P tax	26.8	18.29	86.29	1.83	1.47	0.31	14.65

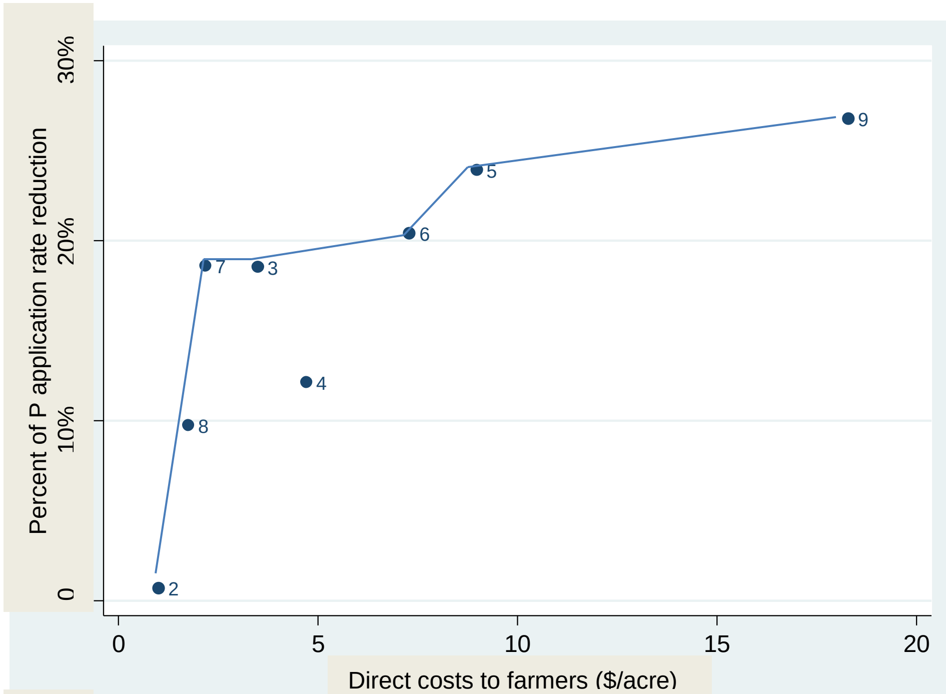
Table 22. The Costs and Cost-Effectiveness of Nutrient Management Policies at Field Level

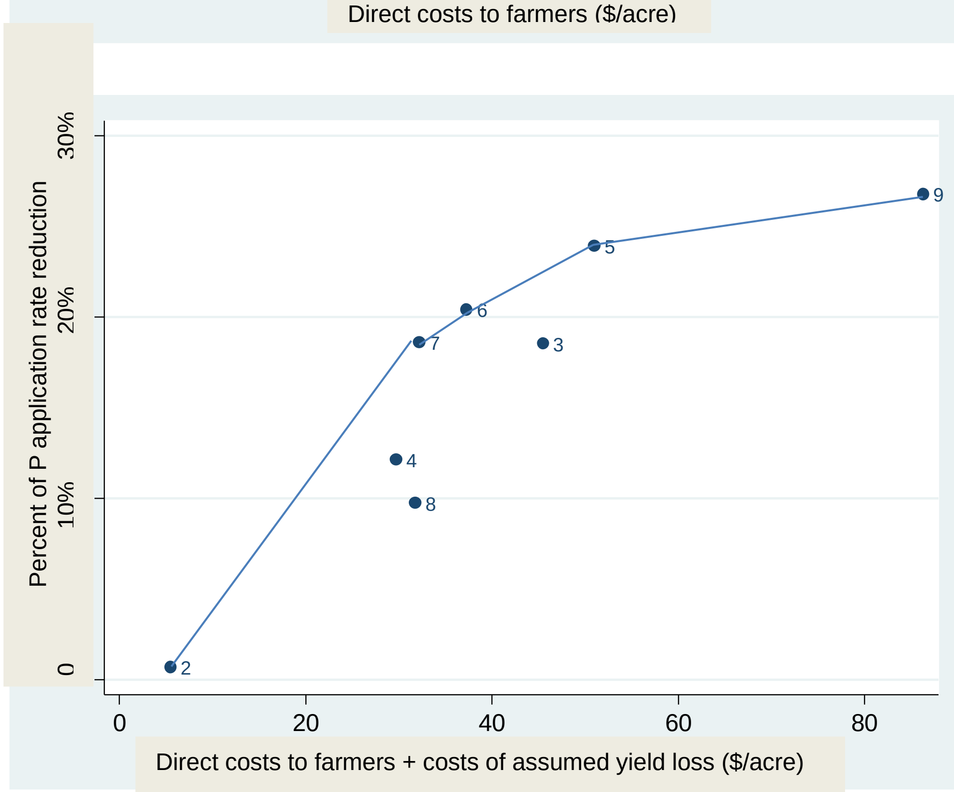
Note: Column (b) "*direct cost to farmers*" denotes event registration fees for 4R education policy, and tax payments plus profit loss due to P limit or tax policies. Column (c) adds cost of assumed potential yield loss, which is 1% yield loss for 4R education policy, 3% for 25% P tax policy, 5% for parcels under 50% P tax and 100 lbs/acre P limit policies, and 10% for 100% P tax policy. Column (d) presents the net policy cost from the social welfare viewpoint, in which tax payments themselves are not treated as net policy costs but costs for administration, monitoring and enforcements are included. Columns (e)-(g) are ratios of percentage reductions in average P application rates per $ of cost, using columns (b)-(d) as costs in the denominator respectively.

“hot spot” subbasins. However, the differences in phosphorus application reductions between spatially-targeted taxes and uniform taxes are not substantial, suggesting possible gains in cost-effectiveness through spatial targeting.

Figure 8 only examines the potential environmental benefits at field level resulting from these alternative policies and ignores the costs associated with these policies. To accurately evaluate the cost-effectiveness of alternative nutrient management policies, it is important to quantify the costs associated with these policies. From a farmer’s perspective, these costs could include the fertilizer taxes they have to pay or registration fees to attend educational events, in addition to the possible reduction in profits resulting from possible yield losses. From a social planner’s perspective, the tax payments from farmers will become revenues for government agencies and thus are just a transfer of wealth. In contrast, a policy would incur net policy costs in terms of its design, implementation, administration, monitoring, and enforcement, which could be viewed as deadweight losses.

Table 22 presents results on the impacts on farmer welfare, deadweight losses, and environmental benefits under these alternative nutrient management scenarios. Given the two different perspectives on costs, for farmers and for society as a whole, I represent the costs in three different ways in table 22. First, column (b) represents the average direct costs for farmers, including which could be registration fees for educational events for the educational campaign policy, and average taxes paid by farmers plus potential profit losses for the fertilizer taxes or limit policies. As shown in table 30 in the Appendix B, I

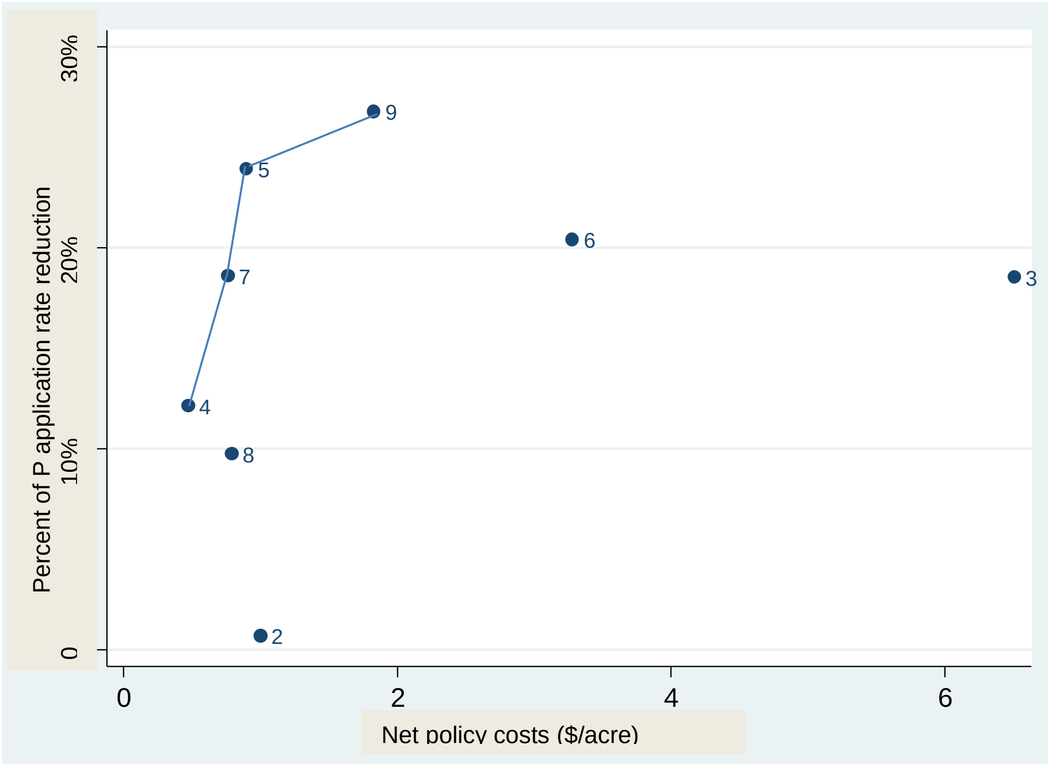




Continued

Figure 9. The Trade-off between Costs and Phosphorus Reduction at Field Level Under Alternative Nutrient Management Policies

Figure 9 continued



Note: The numbers in the x-axis represents different nutrient management policies as shown in Table 21, and they represent: (2) education campaign, (3) phosphorus limit, (4) 25% uniform P tax, (5) 50% uniform P tax, (6) 50% targeted P tax on steeped slope land only, (7) 50% targeted P tax on fields in high runoff pollution subbasin only, (8) 50% targeted P tax on fields with good soil quality only, and (9) 100% uniform P tax, respectively.

did not find statistical evidence that there are yield drags due to changes in phosphorus demand, which is contingent for the range of the prices we analyzed from a single-year cross-sectional data. However, this may stems from the nature of data we use and is contingent on the price ranges we analyzed. And conversations with agronomists reveal that farmers have long-held belief that there will be negative yield drags due to reduction in nutrient applications. As a result, in column (c) I assume there could be negative yield impacts from these agri-environmental policies, and thus the average costs to farmers become the average direct costs shown in column (b) plus average costs of these assumed yield loss. In particular, I assume 1% yield loss for the 4R education policy, 3% for 25% P tax policy, 5% for parcels subject to 50% uniform or targeted P tax and 100 lbs/acre P limit policies, and 10% yield loss for 100% P tax policy. The physical changes in yield loss are then turned into costs by using the prevailing crop prices in 2013.

As discussed earlier, table 22 columns (b) and (c) represent the costs in terms of changes in farmer welfare, while column (d) are the average costs in terms of net policy costs or deadweight losses from a social planner's perspective. These net policy costs incurred for the implementation, monitoring and enforcement of the policies. For example, spatially-targeted policies would result in a much higher monitoring and enforcement costs compared to other policy instruments like uniform tax policies. In a numerical simulation, Lankoski, et al. (2010) find that the enforcement costs for a spatially differentiated fertilizer tax with random monitoring would cost 25-30% of the tax revenue collected. Previous studies on the fertilizer tax policy implementations in Europe also show that the regular implementation would cost 7-10% of the tax revenue collected. As a result, I assume a 35-45% of tax revenue collected as net policy costs for spatially-targeted tax policies, while only assume a 10% implementation costs for uniform tax policies. The phosphorous limit policy is assumed to incur an even greater monitoring and enforcement costs as it is even more difficult to observe and enforce the application rate limit in practice. The costs shown in columns (b) – (d) are average costs for all farmers even when some of them are not in high-runoff-potential subbasins, are not affected by a policy such as targeted taxes based on "hot spot" subbasin.

In this welfare analysis of alternative nutrient management policies, I represent the benefits of the policy as the mean reduction in phosphorus application rates, and they are quantified as the average percentage reduction in phosphorus application rates compared to the baseline scenario. Using the average costs shown in columns (b) – (d) respectively, I compute three ratio-based measures for cost-effectiveness of each policy that are percentage reduction in phosphorus application rates per dollar of average cost, for farmers or for society as a whole. While each of the three ratios has the same property that the higher the ratio is, the more cost-effective the policy, they have different implications on farmer welfare vs. societal deadweight loss. In particular, the first two ratios shown in columns (e) and (f) are concerned with farmer welfare, while the last ratio in column (g) focuses on overall social welfare.

The three ratios shown in table 22 columns (e) to (g) illustrate the cost-effectiveness of these nutrient management policies. A comparison among the three targeting strategies reveals that target parcels based on the location, similar as zonal tax, outperforms spatial targeting that is based on land characteristics. The superiority of subbasin-based targeting is consistent regardless average costs for farmers or average net policy costs are used. Based on the ratio of percent of phosphorus reduction per dollar of the average direct farmer cost, the spatially-targeted fertilizer tax for parcels located in high runoff potential subbasins is most cost-effective policy: a one-dollar increase in the per acre average direct policy cost would lead to an 8% reduction in average application rates.

By contrast, voluntary approaches such as the education campaign are the lowest in both the total reduction in phosphorus and the cost-effectiveness. The 100% uniform tax would lead to the most phosphorus reduction, however, this is achieved through higher tax payments from farmers, which makes it inferior compared to other policies such as 50% uniform tax and 50% targeted tax. Farmers on average have to pay a higher tax for the uniform tax policy as opposed to the targeted policy because every farmer has to pay under the uniform policy scenario while only farmers with certain field or locational characteristics are affected by the spatially targeted policy.

A comparison between column (e) and column (f) shows that, after accounting for potential yield losses into average farmer costs, the advantages of spatially-targeted fertilizer tax policies over uniform tax policies are greatly diminished. This is because the costs of yield loss become the bulk of the total costs to farmers and the cost-savings from spatial-targeting are no longer evident. In addition, the much larger farmer costs reduced the cost-effectiveness ratios significantly: assuming the costs of assumed yield loss is correct; the 50% subbasin-targeted fertilizer tax would only reduce 0.6% of phosphorus application per dollar cost paid by farmers.

Focusing the net policy costs rather than costs to farmers, table 22 column (g) reveals a drastically different picture from the other two ratios shown in columns (e) and (f). The uniform fertilizer tax policies outperform the spatially-targeted policies due to lower monitoring and enforcement costs. For example, a 50% uniform phosphorus tax could lead to a 25.80% reduction in average phosphorus application rate for one dollar net policy cost incurred, but the direct costs to farmers would be 10 times more than the net policy costs.

Figure 9 plots the information shown in table 22, in which the x-axis denotes the costs for farmers or net policy costs, and the y-axis denotes the percentage of phosphorus input rate reductions. The blue lines show the trade-off frontier, in which the policies under this line are dominated by the policies that are on this frontier. These three figures graphically illustrate the dominance of certain agri-environmental policies over another alternative policy instrument, and they reveal different patterns. For example, the 25% uniform tax was dominated by other policies in the first graph when only direct costs to farmers are considered; however, it becomes the most cost-effective on the frontier when the net policy costs are the focus as shown in the third figure. The difference stems from the simplicity of uniform tax policies to monitor, administer and enforce: it takes much more efforts and costs to monitor and enforce the spatially-targeted tax policies or phosphorus application limit policy at the field scale. Another thing to note is that the 50% uniform phosphorus tax and 50% spatially targeted tax based on high-runoff-potential subbasins are on the frontiers for all the three graphs, revealing a consistent and robust performance that might be of interest to policymakers.

There is another thing worth considering in the welfare analysis: the revenues from these policies. And intuitively the uniform tax policies would generate more net revenues compared to spatially-targeted policies due to savings in monitoring and enforcement costs. These net revenues could be used to contribute to the conservation investment budget, which could then pay farmers located in high-runoff-potential subbasins or plotting steeper-sloped fields to encourage their adoption of appropriate and effective best management practices. This stick-n-carrot approach not only helps meeting the policy goal of 40% reduction in phosphorus loadings from Maumee to Lake Erie as table 22 column (a) also reveals that a 50% uniform phosphorus tax could only lead to a 24% reduction in phosphorus application rates, returning the taxes paid by farmers back to the agricultural community also enhances the political feasibility of the policy portfolio. However, the welfare analysis of this policy mix requires at least a model of BMP adoptions and thus is beyond the scope of this chapter.

All these above policy analysis are based on the 1,551 selected agricultural fields, while the Maumee River watershed has over 100,000 agricultural fields that cover at least 3 million acres of row cropland. As a result, it is informative to contemplate the welfare effects of these proposed nutrient management policies at the watershed scale. To do that, I compare the distribution of farms and acreage by state and farm acreage groups between data from our farmer survey and the microdata from 2007 Census of Agriculture. This comparison, shown in table 33 in the Appendix B, reveals that the farmer survey I use is disproportionally concentrated in large farms: my data, which represents about 10% of farms in the watershed, accounts for more than 20% of the total acreage. This means that there are many small-scale farms in the watershed, which could dramatically increase the needs and costs for monitoring and enforcement, especially for a spatially-targeted fertilizer tax. In that context, a policy that is easier to implement and monitor, such as a 50% uniform tax, could present as an attractive policy instrument, as shown in table 22 column (g).

Another caveat of my welfare analysis shown in previous sections is that the environmental benefits are represented by the reductions in phosphorus fertilizer application rates. From the policymakers' perspectives, it is more informative to model the environmental impacts in terms of changes in the size and likelihood of harmful algal blooms in Lake Erie, or at least total and dissolved phosphorus loadings from the entire Maumee River watershed. Current collaboration is also ongoing in linking the predictions from the structural model of farmer behavior to agricultural nutrient loadings using a watershed hydrological model – Soil and Water Assessment Tool (SWAT) that is previously calibrated for western Lake Erie region (Gebremariam, et al. 2014).

Conclusion

Agricultural nutrient pollution has inflicted substantial damages to vital ecosystems both in the United States and worldwide, resulting in a reduction of multiple deliverable ecosystem services, as evidenced by the unprecedented harmful algal blooms in Lake Erie and Gulf of Mexico. Using a farmer survey of 1,551 farmer respondents in western Lake Erie basin, this study develops a structural econometric model of crop and input demand decisions and evaluate the trade-offs between agricultural profits and water quality under alternative nutrient management policy scenarios. By use of a structural model, I am able to quantify the social welfare implications, in terms of changes in both farmer profits and environmental benefits, of alternative nutrient management policies, including non-marginal policies like a 100% fertilizer tax. The model improves on previous studies by not only explicitly incorporating both field-level land and farmer characteristics in the structural model estimation, but also by analyzing the heterogeneity in farmers' elasticity of fertilizer demand due to different environmental attitudes or land characteristics. The mean price elasticities of phosphorus fertilizer demand varies from -0.2 to -0.6, which is within the range of most previous estimates.

The main results also provide evidence on heterogeneity in phosphorus price responsiveness – farmers more familiar with environmental stewardship or growing crops on high soil quality fields have a higher elasticity of fertilizer demand – and the heterogeneity is more evident for corn fields and fields with multi-year fertilization schedules. However, despite the statistically significant coefficients, the magnitude of the differences due to the heterogeneity in fertilizer demand is not very large, which may suggests that the heterogeneous responses among farmers are not as important for the design and implementation of optimal nutrient management policies, at least in my context.

Three measures of the cost-effectiveness of policies are developed: two ratios focus on the percentage reduction in phosphorus application given costs to farmers, while one ratio examine the percentage reduction relative to the net policy costs from the societal welfare perspective. These measures from the perspectives of farmer welfare vs. overall social welfare yield different policy implications: based on ratios that rely on costs to farmers, spatial targeting, especially spatially-targeted phosphorus tax based on location in high-runoff-potential subbasins, could dramatically improve the cost-effectiveness of agri-environmental policies. However, uniform phosphorus tax policies outperform their spatial-targeted counterparts in the cost-effectiveness measure based on net policy cost due to the simplicity and cost-savings in monitoring and enforcement. In addition, regardless of how cost-effectiveness of policy is measured, the 50% uniform phosphorus tax and the 50% spatially-targeted phosphorus tax based on ecologically sensitive subbasins outperform other agri-environmental policies analyzed here.

The main results also show that neither a fertilizer tax nor an education campaign could alone achieve the policy goal of 40% reduction in phosphorus runoff into Lake Erie, although a uniform 50% fertilizer tax could lead to a 24% reduction in mean phosphorus application rates. It is highly likely that other measures such as alternative policy instrument or technological fixes are necessary. For example, a stick-n-carrot approach of combining a uniform fertilizer tax and payments for BMP adoption could be beneficial. Other technological measures such as controlling the runoff as it leaves the field using filter strips, trapping the runoff using in stream wetlands, or even chemical treatment of phosphorus and algal toxins in the lake may all be necessary. And the optimal policy mix depends on the farmers' adoption, their effectiveness in reducing nutrient runoff, political feasibility, and the costs of monitoring and enforcement.

Finally, following the suggestion of Timmins and Schlenker (2009), I complement the estimation of the structural model with the identification of one key parameter – the mean price elasticity of phosphorus fertilizer demand, using a reduced-form panel data model and responses to two hypothetical questions on alternative phosphorus price scenarios. Figure 13 in the Appendix B shows the distribution of hypothetical application rates given the alternative, hypothetical phosphorus fertilizer prices. The ranges of hypothetical prices are based on actual historical prices that farmers have experienced over the last 15 years, which helps mitigate the hypothetical bias and induces true responses. The results show the reduced-form panel data model based on these two hypothetical questions help restrict the estimated elasticity within a more reasonable range. The results suggest that incorporating simple hypothetical scenarios into the survey design of cross-sectional data could present as a useful and low-cost complement to standard modeling approach and help the identification of the key parameters, like the price elasticity of fertilizer demand analyzed here.

References

Abbott, Joshua K., and H. Allen Klaiber. 2011. "An Embarrassment of Riches: Confronting Omitted Variable Bias and Multi-Scale Capitalization in Hedonic Price Models." *Review of Economics and Statistics* 93:1331-1342.

Anderson, Soren T., and Sarah E. West. 2006. "Open Space, Residential Property Values, and Spatial Context." *Regional Science and Urban Economics* 36:773-789.

Ando, Amy W., and Mindy L. Mallory. 2012. "Optimal Portfolio Design to Reduce Climate-Related Conservation Uncertainty in the Prairie Pothole Region." *Proceedings of the National Academy of Sciences* 109:6484-6489.

Anselin, Luc, and Sheri Hudak. 1992. "Spatial Econometrics in Practice: A Review of Software Options." *Regional Science and Urban Economics* 22:509-536.

Arnade, Carlos, and David Kelch. 2007. "Estimation of Area Elasticities from a Standard Profit Function." *American Journal of Agricultural Economics* 89:727-737.

Ayyagari, Padmaja, et al. 2013. "Understanding Heterogeneity in Price Elasticities in the Demand for Alcohol for Older Individuals." *Health Economics* 22:89-105.

Babcock, Bruce A., et al. 1997. "Targeting Tools for the Purchase of Environmental Amenities." *Land Economics*:325-339.

Bajari, Patrick, et al. 2012. "A Rational Expectations Approach to Hedonic Price Regressions with Time-Varying Unobserved Product Attributes: The Price of Pollution." *American Economic Review* 102:1898-1926.

Bastian, Chris T., et al. 2002. "Environmental Amenities and Agricultural Land Values: A Hedonic Model Using Geographic Information Systems Data." *Ecological Economics* 40:337-349.

Baum, Christopher F., Mark E. Schaffer, and Steven Stillman. 2003. "Instrumental Variables and Gmm: Estimation and Testing." *Stata journal* 3:1-31.

Bayer, Patrick, and Christopher Timmins. 2007. "Estimating Equilibrium Models of Sorting across Locations*." *The Economic Journal* 117:353-374.

Bento, Antonio, Charles Towe, and Jacqueline Geoghegan. 2007. "The Effects of Moratoria on Residential Development: Evidence from a Matching Approach." *American Journal of Agricultural Economics* 89:1211-1218.

Binswanger, Hans P. 1974. "A Cost Function Approach to the Measurement of Elasticities of Factor Demand and Elasticities of Substitution." *American Journal of Agricultural Economics*:377-386.

Bishop, Kelly C., and Christopher Timmins. "Hedonic Prices and Implicit Markets: Estimating Marginal Willingness to Pay for Differentiated Products without Instrumental Variables." National Bureau of Economic Research, 2011.

Blomendahl, Ben H., Richard K. Perrin, and Bruce B. Johnson. 2011. "The Impact of Ethanol Plants on Surrounding Farmland Values: A Case Study." *Land Economics* 87:223-232.

Busso, Matias, Jesse Gregory, and Patrick Kline. 2013. "Assessing the Incidence and Efficiency of a Prominent Place Based Policy." *The American Economic Review* 103:897-947.

Caliendo, Marco, and Sabine Kopeinig. 2008. "Some Practical Guidance for the Implementation of Propensity Score Matching." *Journal of Economic Surveys* 22:31-72.

Capozza, Dennis R., and Robert W. Helsley. 1989. "The Fundamentals of Land Prices and Urban Growth." *Journal of Urban Economics* 26:295-306.

Cappiello, Dina, and Matt Apuzzo 2013 "The Secret Environmental Cost of U.S. Ethanol Policy." In *Associated Press, November 12, 2013.*

Carman, Hoy F. 1979. "The Demand for Nitrogen, Phosphorous and Potash Fertilizer Nutrients in the Western United States." *Western Journal of Agricultural Economics*:23-31.

Cavailhès, Jean, and Pierre Wavresky. 2003. "Urban Influences on Periurban Farmland Prices." *European Review of Agricultural Economics* 30:333-357.

Chambers, Robert G. 1988. *Applied Production Analysis: A Dual Approach*: Cambridge University Press.

Chetty, Raj. "Sufficient Statistics for Welfare Analysis: A Bridge between Structural and Reduced-Form Methods." National Bureau of Economic Research, 2008.

Chicoine, David L. 1981. "Farmland Values at the Urban Fringe: An Analysis of Sale Prices." *Land Economics* 57:353-362.

Claassen, Roger, and Richard D. Horan. 2001. "Uniform and Non-Uniform Second-Best Input Taxes." *Environmental and Resource Economics* 19:1-22.

Cohen, Jeffrey P., Cletus C. Coughlin, and David A. Lopez. 2012. "The Boom and Bust of Us Housing Prices from Various Geographic Perspectives." *Federal Reserve Bank of St. Louis Review* 94.

Crump, Richard K., et al. 2009. "Dealing with Limited Overlap in Estimation of Average Treatment Effects." *Biometrika* 96:187-199.

d'Aspremont, Claude, J. Jaskold Gabszewicz, and J.-F. Thisse. 1979. "On Hotelling's" Stability in Competition"." *Econometrica* 47:1145-1150.

De Vor, Friso, and Henri LF De Groot. 2011. "The Impact of Industrial Sites on Residential Property Values: A Hedonic Pricing Analysis from the Netherlands." *Regional Studies* 45:609-623.

Denbaly, Mark, and Harry Vroomen. 1993. "Dynamic Fertilizer Nutrient Demands for Corn: A Cointegrated and Error-Correcting System." *American Journal of Agricultural Economics* 75:203-209.

Diamond, Alexis, and Jasjeet S. Sekhon. 2013. "Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies." *Review of Economics and Statistics* 95:932-945.

Dillman, Don A. 2011. *Mail and Internet Surveys: The Tailored Design Method--2007 Update with New Internet, Visual, and Mixed-Mode Guide*: John Wiley & Sons.

Dobos, R., H. Sinclair, and K. Hipple. 2008. "User Guide National Commodity Crop Productivity Index (Nccpi) Version 1.0." *US Department of Agriculture, Natural Resources Conservation Service*.

Fackler, Paul L., and Barry K. Goodwin. 2001. "Spatial Price Analysis." *Handbook of Agricultural Economics* 1:971-1024.

Farm Net Services 2012 "Ohio Grain Elevators." In *Farm Equipment and Agriculture Business Directory, http://www.farmnetservices.com/farm/Grain%20Elevators/OHIO%20GRAIN%20ELEVATORS/66-0.html*.

Featherstone, Allen M., and Barry K. Goodwin. 1993. "Factors Influencing a Farmer's Decision to Invest in Long-Term Conservation Improvements." *Land Economics*:67-81.

Gallagher, Paul W. 2006. "Energy Production with Biomass: What Are the Prospects?" *Choices: The Magazine of Food, Farm, and Resource Issues* 21:21-25.

Gebremariam, Seyoum Y., et al. 2014. "A Comprehensive Approach to Evaluating Watershed Models for Predicting River Flow Regimes Critical to Downstream Ecosystem Services." *Environmental Modelling & Software* 61:121-134.

Gillingham, Kenneth. 2014. "Identifying the Elasticity of Driving: Evidence from a Gasoline Price Shock in California." *Regional Science and Urban Economics* 47:13-24.

Gloy, Brent A., et al. 2011 "Farmland Values: Current and Future Prospects." In *Center for Commercial Agriculture Staff Paper, Department of Agricultural Economics, Purdue University*. West Lafayette, IN.

Goetz, Renan U., and David Zilberman. 2000. "The Dynamics of Spatial Pollution: The Case of Phosphorus Runoff from Agricultural Land." *Journal of Economic Dynamics and Control* 24:143-163.

Griliches, Zvi. 1958. "The Demand for Fertilizer: An Economic Interpretation of a Technical Change." *Journal of Farm Economics* 40:591-606.

---. 1959. "The Demand for Inputs in Agriculture and a Derived Supply Elasticity." *Journal of Farm Economics* 41:309-322.

Guiling, Pam, B. Wade Brorsen, and Damona Doye. 2009. "Effect of Urban Proximity on Agricultural Land Values." *Land Economics* 85:252-264.

Gunjal, Kisan R., Roland K. Roberts, and Earl O. Heady. 1980. "Fertilizer Demand Functions for Five Crops in the United States." *Southern Journal of Agricultural Economics* 12:111-116.

Hansen, Lars Gårn. 2004. "Nitrogen Fertilizer Demand from Danish Crop Farms: Regulatory Implications of Heterogeneity." *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie* 52:313-331.

Hansen, Lars Peter. 1982. "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica: Journal of the Econometric Society*:1029-1054.

Hansen, Line Block, and Lars Gårn Hansen. 2014. "Can Non-Point Phosphorus Emissions from Agriculture Be Regulated Efficiently Using Input-Output Taxes?" *Environmental and Resource Economics* 58:109-125.

Hardie, Ian W., Tulika A. Narayan, and Bruce L. Gardner. 2001. "The Joint Influence of Agricultural and Nonfarm Factors on Real Estate Values: An Application to the Mid-Atlantic Region." *American Journal of Agricultural Economics* 83:120-132.

Harding, Matthew, and Michael Lovenheim. "The Effect of Prices on Nutrition: Comparing the Impact of Product-and Nutrient-Specific Taxes." National Bureau of Economic Research, 2014.

Heady, Earl O., and Martin H. Yeh. 1959. "National and Regional Demand Functions for Fertilizer." *Journal of Farm Economics*:332-348.

Heckman, James J., Hidehiko Ichimura, and Petra Todd. 1998. "Matching as an Econometric Evaluation Estimator." *The Review of Economic Studies* 65:261-294.

Heckman, James, and Salvador Navarro-Lozano. 2004. "Using Matching, Instrumental Variables, and Control Functions to Estimate Economic Choice Models." *Review of Economics and Statistics* 86:30-57.

Henderson, Jason, and Brent A. Gloy. 2009. "The Impact of Ethanol Plants on Cropland Values in the Great Plains." *Agricultural Finance Review* 69:36-48.

Henderson, Jason, and Sean Moore. 2006. "The Capitalization of Wildlife Recreation Income into Farmland Values." *Journal of Agricultural and Applied Economics* 38:597-620.

Hendricks, Nathan P., et al. 2014. "The Environmental Effects of Crop Price Increases: Nitrogen Losses in the U.S. Corn Belt." *Journal of Environmental Economics and Management* 68:507-526.

Hite, Diane, et al. 2001. "Property-Value Impacts of an Environmental Disamenity: The Case of Landfills." *The Journal of Real Estate Finance and Economics* 22:185-202.

Howard, Gregory, and Brian E. Roe. 2013. "Stripping Because You Want to Versus Stripping Because the Money Is Good: A Latent Class Analysis of Farmer Preferences Regarding Filter Strip Programs." *AAEA and CAES Joint Annual Meeting, Washington, DC August*.

Huang, Haixiao, et al. 2006. "Factors Influencing Illinois Farmland Values." *American Journal of Agricultural Economics* 88:458-470.

Iho, Antti, and Marita Laukkanen. 2012. "Precision Phosphorus Management and Agricultural Phosphorus Loading." *Ecological Economics* 77:91-102.

Imbens, Guido W., and Jeffrey M. Wooldridge. 2009. "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature* 47:5-86.

Jacobsen, Mark R. 2013. "Evaluating Us Fuel Economy Standards in a Model with Producer and Household Heterogeneity." *American Economic Journal: Economic Policy* 5:148-187.

Kleibergen, Frank, and Richard Paap. 2006. "Generalized Reduced Rank Tests Using the Singular Value Decomposition." *Journal of Econometrics* 133:97-126.

Kohlhase, Janet E. 1991. "The Impact of Toxic Waste Sites on Housing Values." *Journal of Urban Economics* 30:1-26.

Kuminoff, Nicolai V., and Jaren C. Pope. 2013. "The Value of Residential Land and Structures During the Great Housing Boom and Bust." *Land Economics* 89:1-29.

Kurkalova, Lyubov, Catherine Kling, and Jinhua Zhao. 2006. "Green Subsidies in Agriculture: Estimating the Adoption Costs of Conservation Tillage from Observed Behavior." *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie* 54:247-267.

Lambert, Dayton M., et al. 2008. "Ethanol Plant Location Determinants and County Comparative Advantage." *Journal of Agricultural and Applied Economics* 40:117-135.

Lankoski, Jussi, Erik Lichtenberg, and Markku Ollikainen. 2010. "Agri-Environmental Program Compliance in a Heterogeneous Landscape." *Environmental and Resource Economics* 47:1-22.

Lankoski, Jussi, and Markku Ollikainen. 2003. "Agri-Environmental Externalities: A Framework for Designing Targeted Policies." *European Review of Agricultural Economics* 30:51-75.

Laukkanen, Marita, and Céline Nauges. 2014. "Evaluating Greening Farm Policies: A Structural Model for Assessing Agri-Environmental Subsidies." *Land Economics* 90:458-481.

Lee, Lung-Fei. 1983. "Generalized Econometric Models with Selectivity." *Econometrica: Journal of the Econometric Society*:507-512.

Lincoln Institute of Land Policy 2012 "Land and Property Values in the U.S." In http://www.lincolninst.edu/subcenters/land-values/metro-area-land-prices.asp ed.

Liu, Weiwei. 2014. "Modeling Gasoline Demand in the United States: A Flexible Semiparametric Approach." *Energy Economics* 45:244-253.

Livanis, Grigorios, et al. 2006. "Urban Sprawl and Farmland Prices." *American Journal of Agricultural Economics* 88:915-929.

Low, Sarah A., and Andrew M. Isserman. 2009. "Ethanol and the Local Economy Industry Trends, Location Factors, Economic Impacts, and Risks." *Economic Development Quarterly* 23:71-88.

Lynch, Lori, Wayne Gray, and Jacqueline Geoghegan. 2007. "Are Farmland Preservation Program Easement Restrictions Capitalized into Farmland Prices? What Can a Propensity Score Matching Analysis Tell Us?" *Applied Economic Perspectives and Policy* 29:502-509.

Ma, Shan, and Scott M. Swinton. 2011. "Valuation of Ecosystem Services from Rural Landscapes Using Agricultural Land Prices." *Ecological Economics* 70:1649-1659.

Maddala, Gangadharrao Soundalyarao. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*: Cambridge University Press.

McNew, Kevin, and Duane Griffith. 2005. "Measuring the Impact of Ethanol Plants on Local Grain Prices." *Applied Economic Perspectives and Policy* 27:164-180.

Meier, Petra Sylvia, Robin Purshouse, and Alan Brennan. 2010. "Policy Options for Alcohol Price Regulation: The Importance of Modelling Population Heterogeneity." *Addiction* 105:383-393.

Michalak, Anna M., et al. 2013. "Record-Setting Algal Bloom in Lake Erie Caused by Agricultural and Meteorological Trends Consistent with Expected Future Conditions." *Proceedings of the National Academy of Sciences* 110:6448-6452.

Moschini, Giancarlo. 1988. "The Cost Structure of Ontario Dairy Farms: A Microeconometric Analysis." *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie* 36:187-206.
---. 1988. "A Model of Production with Supply Management for the Canadian Agricultural Sector." *American Journal of Agricultural Economics* 70:318-329.
Nehring, Richard, et al. 2006. "Urban Influence on Costs of Production in the Corn Belt." *American Journal of Agricultural Economics* 88:930-946.
Newburn, David A., Peter Berck, and Adina M. Merenlender. 2006. "Habitat and Open Space at Risk of Land-Use Conversion: Targeting Strategies for Land Conservation." *American Journal of Agricultural Economics* 88:28-42.
Nickerson, Cynthia J., and Lori Lynch. 2001. "The Effect of Farmland Preservation Programs on Farmland Prices." *American Journal of Agricultural Economics* 83:341-351.
Nickerson, Cynthia J., et al. 2012. *Trends in Us Farmland Values and Ownership*: US Department of Agriculture, Economic Research Service.
Nickerson, Cynthia, and Wendong Zhang 2014 "Modeling the Determinants of Farmland Values in the United States." In Joshua Duke, and JunJie Wu eds. *The Oxford Handbook of Land Economics.* Oxford University Press, pp. 111-138.
Norris, Patricia E., and Sandra S. Batie. 1987. "Virginia Farmers' Soil Conservation Decisions: An Application of Tobit Analysis." *Southern Journal of Agricultural Economics* 19:79-90.
Ohio Department of Agriculture 2012 "Ohio Licensed Commodity Handlers." In *Grain Warehouse Program.*
Ohio Department of Transportation 2012 "Active Rails, Airports, Cities and Townships." In *ESRI/Statewide Digital Data.*
Ohio Environmental Protection Agency. 2013. *Ohio Lake Erie Phosphorus Task Force Ii Final Report*.
Ohio Ethanol Council 2012 "Ohio Ethanol Industry Quick Facts." In *http://ohethanol.com/ohio_ethanol*.
Palmquist, Raymond B. 1989. "Land as a Differentiated Factor of Production: A Hedonic Model and Its Implications for Welfare Measurement." *Land Economics* 65:23-28.
---. 2005. "Property Value Models." *Handbook of Environmental Economics* 2:763-819.
Palmquist, Raymond B., and Leon E. Danielson. 1989. "A Hedonic Study of the Effects of Erosion Control and Drainage on Farmland Values." *American Journal of Agricultural Economics* 71:55-62.
Partridge, Mark D., et al. 2008. "Lost in Space: Population Growth in the American Hinterlands and Small Cities." *Journal of Economic Geography*:lbn038.
Penm, Jammie H., and David P. Vincent. 1987. "Some Estimates of the Price Elasticity of Demand for Phosphatic and Nitrogenous Fertilisers." *Australian Journal of Agricultural Economics* 31:65-73.
Pitt, Mark M. 1983. "Farm-Level Fertilizer Demand in Java: A Meta-Production Function Approach." *American Journal of Agricultural Economics* 65:502-508.

Rabotyagov, Sergey, et al. 2010. "Least-Cost Control of Agricultural Nutrient Contributions to the Gulf of Mexico Hypoxic Zone." *Ecological Applications* 20:1542-1555.

Rabotyagov, Sergey S., et al. 2014. "Cost-Effective Targeting of Conservation Investments to Reduce the Northern Gulf of Mexico Hypoxic Zone." *Proceedings of the National Academy of Sciences* 111:18530-18535.

Rajagopal, Deepak, Gal Hochman, and David Zilberman. 2011. "Indirect Fuel Use Change (Ifuc) and the Lifecycle Environmental Impact of Biofuel Policies." *Energy Policy* 39:228-233.

Reutter, Jeffrey Michael, Lake Erie Millennium Network, and Synthesis Team. 2011. *Lake Erie Nutrient Loading and Harmful Algal Blooms: Research Findings and Management Implications*: Ohio Sea Grant College Program, Ohio State University.

Ricardo, David. 1996. *The Principles of Political Economy, and Taxation*. Amherst, NY: Prometheus Books [originally published in 1817].

Ricker-Gilbert, Jacob, Thomas S. Jayne, and Ephraim Chirwa. 2011. "Subsidies and Crowding Out: A Double-Hurdle Model of Fertilizer Demand in Malawi." *American Journal of Agricultural Economics*:aaq122.

Rosen, Sherwin. 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *The Journal of Political Economy* 82:34-55.

Rosenbaum, Paul R., and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70:41-55.

Rubin, Donald B. 1980. "Bias Reduction Using Mahalanobis-Metric Matching." *Biometrika* 36:293-298.

Schnitkey, Gary D., and Bruce J. Sherrick. 2011. "Income and Capitalization Rate Risk in Agricultural Real Estate Markets." *Choices* 26.

Shi, Yue Jin, Timothy T. Phipps, and Dale Colyer. 1997. "Agricultural Land Values under Urbanizing Influences." *Land Economics* 73:90-100.

Shumway, C. Richard, and William P. Alexander. 1988. "Agricultural Product Supplies and Input Demands: Regional Comparisons." *American Journal of Agricultural Economics* 70:153-161.

Shumway, C. Richard, Roberto R. Saez, and Pablo E. Gottret. 1988. "Multiproduct Supply and Input Demand in Us Agriculture." *American Journal of Agricultural Economics* 70:330-337.

Stock, James H., and Motohiro Yogo. 2005. "Testing for Weak Instruments in Linear Iv Regression." *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg*.

Tiffany, Douglas G. 2009. "Economic and Environmental Impacts of Us Corn Ethanol Production and Use." *Regional Economic Development* 5:42-58.

Timmins, Christopher, and Wolfram Schlenker. 2009. "Reduced-Form Versus Structural Modeling in Environmental and Resource Economics." *Annu. Rev. Resour. Econ.* 1:351-380.

Towe, Charles, and Constant I. Tra. 2012. "Vegetable Spirits and Energy Policy." *American Journal of Agricultural Economics* 95:1-16.

U.S. Census TIGER/Line 2012 "Census Block Shapefiles with 2010 Census Population and Housing Unit Counts." In *ftp://ftp2.census.gov/geo/tiger/TIGER2010BLKPOPHU/*.

Villezca-Becerra, Pedro A., and C. Richard Shumway. 1992. "Multiple-Output Production Modeled with Three Functional Forms." *Journal of Agricultural and Resource Economics*:13-28.

Wallander, Steve, Roger Claassen, and Cynthia Nickerson. 2011. "The Ethanol Decade: An Expansion of Us Corn Production, 2000-09." *USDA-ERS Economic Information Bulletin*.

Williamson, James M. 2011. "The Role of Information and Prices in the Nitrogen Fertilizer Management Decision: New Evidence from the Agricultural Resource Management Survey." *Journal of Agricultural and Resource Economics*:552-572.

Wilson, Robyn S., Gregory Howard, and Elizabeth A. Burnett. 2014. "Improving Nutrient Management Practices in Agriculture: The Role of Risk-Based Beliefs in Understanding Farmers' Attitudes toward Taking Additional Action." *Water Resources Research* 50:6735-6746.

Wu, JunJie, et al. 2004. "From Microlevel Decisions to Landscape Changes: An Assessment of Agricultural Conservation Policies." *American Journal of Agricultural Economics* 86:26-41.

Wu, JunJie, and Bruce A. Babcock. 1998. "The Choice of Tillage, Rotation, and Soil Testing Practices: Economic and Environmental Implications." *American Journal of Agricultural Economics* 80:494-511.

Xabadia, Angels, Renan U. Goetz, and David Zilberman. 2008. "The Gains from Differentiated Policies to Control Stock Pollution When Producers Are Heterogeneous." *American Journal of Agricultural Economics* 90:1059-1073.

Zhang, Wendong, and Cynthia J. Nickerson. forthcoming. "The Housing Market Bust and Farmland Values: Identifying the Changing Influence of Proximity to Urban Centers." *Land Economics*.

Zhao, Zhong. 2004. "Using Matching to Estimate Treatment Effects: Data Requirements, Matching Metrics, and Monte Carlo Evidence." *Review of Economics and Statistics* 86:91-107.

Appendix A: Additional Figures and Tables for Chapter 2

Dependent variable	(I) Dist_Ethanol		(II) Dist_Ethanol* Post construction dummy	
	Coef.	Std. Err.	Coef.	Std. Err.
Assessed land value % of total assessed	0.3959	0.3967	-0.3078	0.5078
Total acres	0.0006	0.0028	0.0020	0.0034
Total acres squared	5.13E-06	0.0000	-3.17E-06	0.0000
NCCPI	0.0004***	0.0001	0.0002**	0.0001
Prime farmland	0.7857**	0.3560	0.9595**	0.4387
Steep slope (>15 degrees)	-0.3574	0.3211	0.1460	0.4698
Building area % of parcel	0.5002	0.4272	0.5067	0.6605
Forest area % of parcel	-0.2076	0.7060	0.0401	0.7424
Distance to highway ramp	0.0402	0.0420	-0.0077	0.0468
Distance to nearest city	-0.0217	0.0252	-0.0276	0.0266
Incremental distance to second nearest city	0.0960***	0.0238	0.0397	0.0249
Surrounding population within 25 miles	0.0031*	0.0017	-0.0010	0.0017
Gravity index of three nearest cities	0.0003	0.0010	0.0013	0.0011
Distance to railways	0.0853*	0.0448	0.2016***	0.0512
Distance to nearest grain elevator	0.3628***	0.0393	0.2061***	0.0489
Distance to nearest agricultural terminal	0.1671***	0.0195	0.0819***	0.0220
Capacity-weighted dist to other ethanol plants	0.1701***	0.0267	-0.0752***	0.0260
Capacity-weighted distance to other terminals	0.0002***	0.0000	0.0001***	0.0000
Avg_Dist_Ethanol * Post construction	-7.04E-06*	0.0000	-2.91E-05***	0.0000
Avg_Dist_Terminal * Post construction	-0.0056	0.0079	0.3103***	0.0105
Intercept	-18.5981***	2.5858	-8.1465***	2.8081
Year FE	yes		yes	
County FE	yes		yes	
Adjusted R^2	0.852		0.7707	
Number of observations	3343		3343	

Table 23. First Stage Regressions of the Instrumental Variables Estimation

Capacity-weighted distances to other agricultural terminals * post_dummy	Coef.	Std. Err.
Assessed land value % of total assessed	-0.0534	1.3204
Total acres	-0.0127	0.0093
Total acres squared	4.18E-05	0.0000
National Commodity Crops Productivity Index	-8.39E-05	0.0002
Prime farmland	0.4471	1.0687
Steep slope (>15 degrees)	-0.1979	1.1024
Building area % of parcel	2.3609	1.5945
Forest area % of parcel	3.8843*	2.1543
Wetland area % of parcel	-9.6017	19.6074
Distance to highway ramp	0.1440	0.1251
Distance to nearest city	-0.0639	0.0709
Incremental distance to second nearest city	-0.0374	0.0593
Surrounding population within 25 miles	0.0118***	0.0040

Continued

Table 24. Indirect Test for the Validity of the Instruments

Table 24 continued

Gravity index of three nearest cities	-0.0021	0.0022
Distance to railways	-0.3292**	0.1431
Distance to nearest grain elevator	0.2861**	0.1164
Distance to nearest agricultural terminal	0.1603***	0.0528
Intercept	80.9746***	11.6097
Year FE	yes	
County FE	yes	
Adjusted R^2	0.852	
Number of observations	3343	

(I) Weak identification test	
Kleibergen-Paap rk Wald F statistic	41.005
Cragg-Donald Wald F statistic	68.886
Stock-Yogo weak ID test critical value for 10% maximal IV size	16.87

Continued

Table 25. Tests of Weak Identification, Overidentification of all Instruments and Endogeneity Test of Endogenous Regressors

Table 25 continued

(II) Test of overidentifying restrictions	
Hansen J statistic	0.400
p-value	0.819
(III) Endogeneity test of endogenous regressors	
GMM distance test of endogeneity statistic	5.581
p-value	0.0614

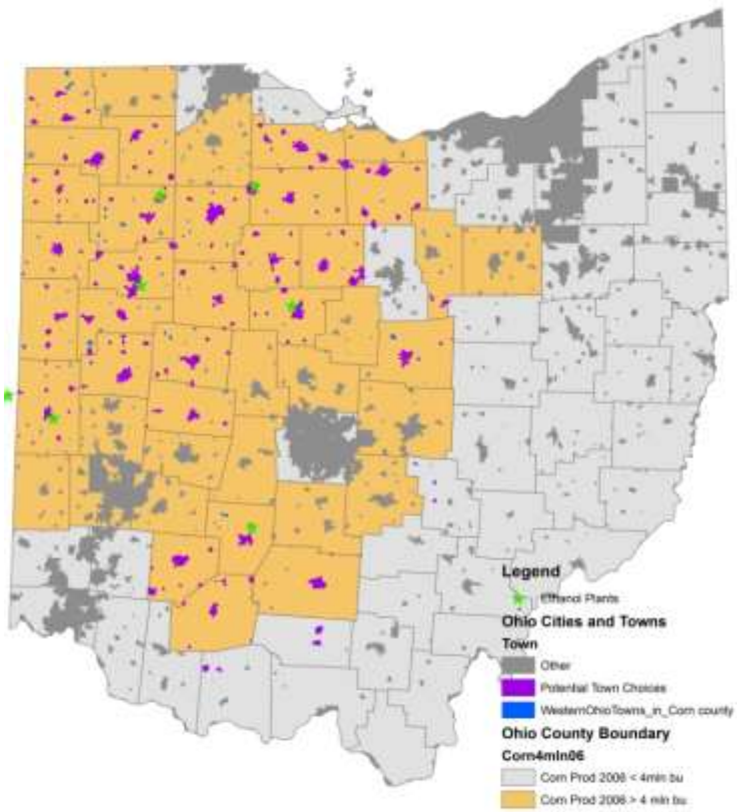
Nominal farmland values ($/acre)	(I)		(II)		(III)	
	Coef.	Robust SE	Coef.	Robust SE	Coef.	Robust SE
Capacity-weighted dist to other ethanol plants	10.27	8.02	14.86	16.04		
Avg_Dist_Ethanol * Post construction	-0.1505	2.77	-15.35***	5.66		
Capacity-weighted distance to other terminals	0.0014	0.0054	-0.0053	0.0102	0.0005	0.0097
Avg_Dist_Terminal * Post construction	-0.0026*	0.0014	0.0008	0.0028	-2.6521	3.6714
Assessed land value % of total assessed	-3771.32***	144.92	-3356.72***	375.05	-3315.51***	373.21
Total acres	-26.04***	1.04	-40.14***	3.07	-40.17***	3.05
Total acres squared	0.013***	0.0017	0.08***	0.01	0.078***	0.013
NCCPI	0.0035	0.026	0.0355	0.048	0.0375	0.047
Prime farmland	-75.02	116.35	-268.38	267.44	-378.34	258.47
Steep slope (>15 degrees)	-112.98*	58.87	319.30	257.55	240.79	247.45
Building area % of parcel	-50.00	268.35	-572.97	394.79	-495.17	396.46
Forest area % of parcel	-4.33	159.63	-398.24	521.17	-326.87	496.90

Continued

Table 26. Regression of Farmland Values on Instruments

Table 26 continued

Nominal farmland values ($/acre)	(I)		(II)		(III)	
	Coef.	Robust SE	Coef.	Robust SE	Coef.	Robust SE
Wetland area % of parcel	-194.69	872.00	798.53	3318.81	1783.28	3318.81
Distance to highway ramp	-38.15**	15.19	-32.62	27.90	-32.10	27.39
Distance to nearest city	-63.66***	8.09	-28.20*	15.84	-24.89*	15.53
Incremental distance to 2nd nearest city	-37.84**	6.09	-32.25**	13.94	-33.54**	13.53
Surrounding population within 25 miles	-0.65**	0.30	-0.80	0.83	-0.04	0.79
Gravity index of three nearest cities	0.0003*	0.0002	0.60	0.56	0.71	0.54
Distance to railways	-3.77	17.36	-2.04	32.27	-4.45	31.74
Distance to nearest grain elevator	-1.39	9.64	-39.54	26.35	-39.43	25.82
Distance to nearest agricultural terminal	-34.70***	5.34	-3.06	11.84	-5.41	11.69
Intercept	14639.57***	2299.78	9336.27***	1736.01	8471.05***	1612.21
Year and County fixed effects	yes		yes		yes	
Adjusted R^2	0.2618		0.2426		0.2358	
Number of observations	16434		3443		3541	



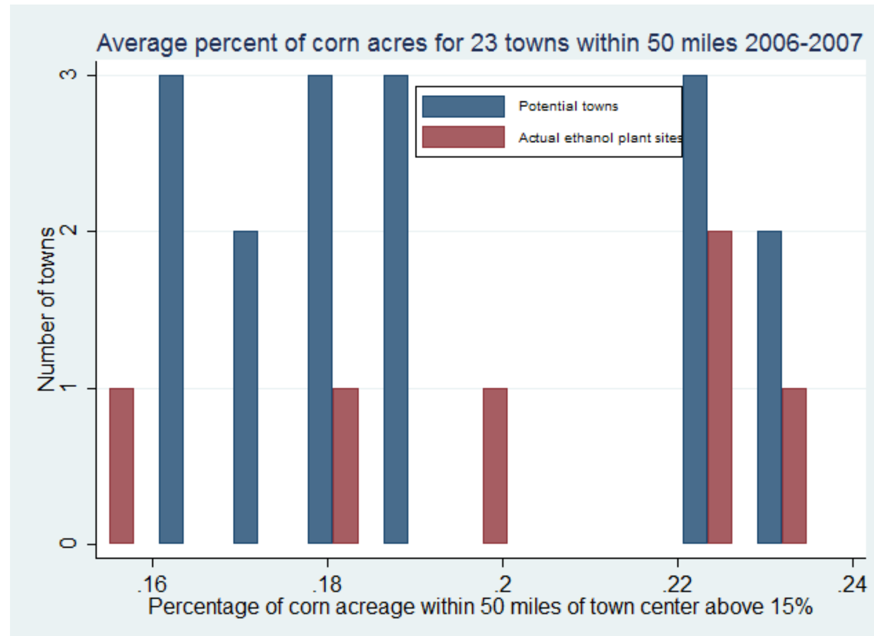


Figure 10. Alternative Towns as Sites for Ethanol Plants and Percentage of Corn Acreage within 50 Miles from Actual ethanol Plant and Candidate Towns

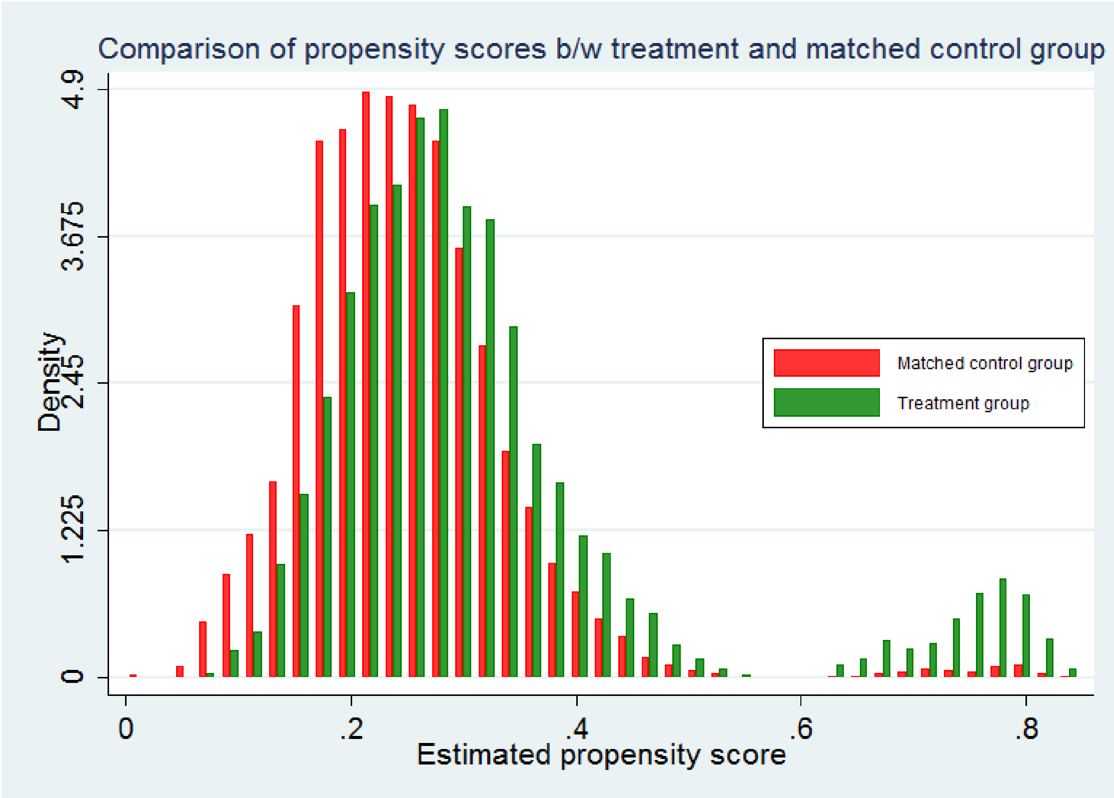


Figure 11. The Comparison of Propensity Score between Treatment and Matched Control Groups for Matching based on Proximity to Ethanol Plants

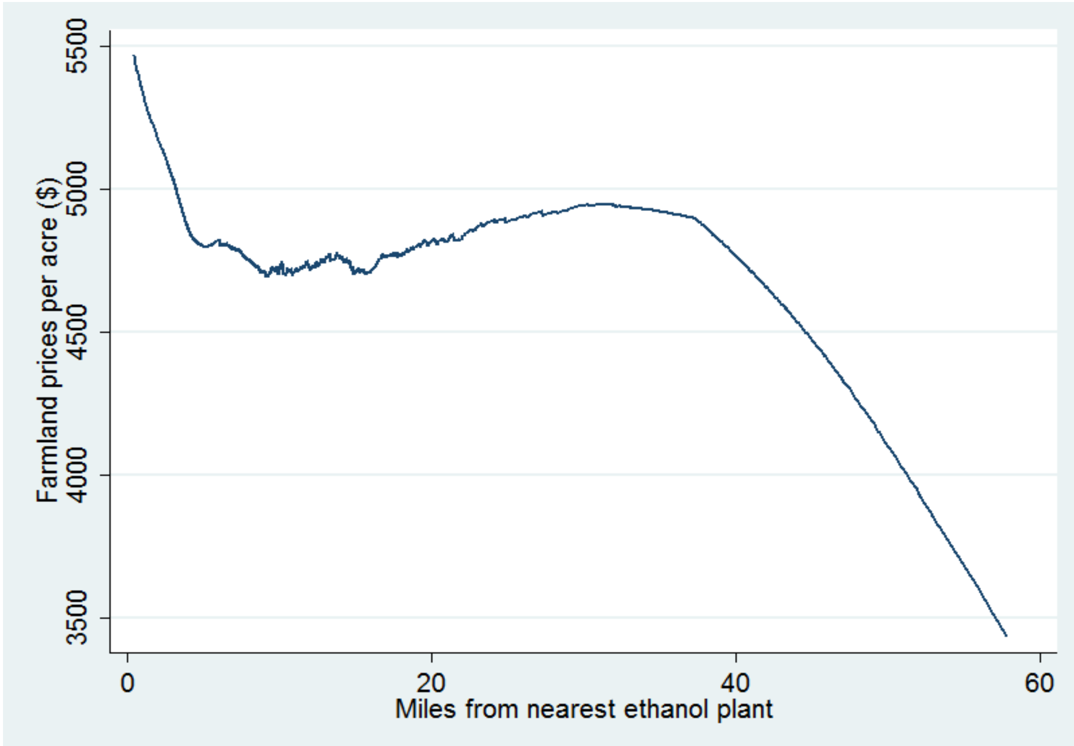


Figure 12. Nonparametric Estimation of Farmland Values with respect to Proximity to Nearest Ethanol Plant

Appendix B: Additional Figures and Tables for Chapter 3

	Coefficient	Std. Err.
Input and Output Prices		
P price	-0.00002	0.00015
N price	-9.03E-06	0.00005
Expected corn price	-0.0001	0.0002
Expected soybean price	0.00006	0.00013
Farmer Characteristics		
Farmilar with 4R	-0.0003	0.0132
More risk loving	0.0071	0.0072
Age	-0.00004	0.00176
Education	0.0072	0.0123
Years of farming experience	-0.0007	0.0011
Female operator	-0.0848	0.2397
Farm income	0.0658***	0.0170
Farm has livestock	0.0052	0.0350
Farm has crop insurance	-0.0255	0.0377
Farm Characteristics		
Farm acres	-0.00008***	0.00003
% good soil for a farm	-0.0004	0.0005
% poor soil for a farm	-0.0007	0.0007
Number of fields	-0.00001	0.0007
Intercept	0.2633	0.1918

Table 27. Regressions on Mix of Crop Production at the Farm Level

	Coefficient	Std. Err.
Phosphorus Fertilizer Prices with Interactions		
Normalized P fertilizer price	-1.967***	0.602
Normalized P price * familiar 4R	0.880***	0.214
Normalized P price * slope	-0.348	0.507
Normalized P price * good soil	-1.116*	0.644
Normalized P price * poor soil	0.331	0.641
Input and Output Prices		
Normalized N price	0.149***	0.045
Normalized expected corn price	-0.129	0.240
Normalized manure price	526.655	3866.177
Farmer Characteristics		
Farmilar with 4R	-55.886***	12.580
More risk loving	-3.277***	1.136
Age	-0.464*	0.282
Education	-2.536	1.965
Years of farming experience	0.247	0.181
Female operator	13.317	32.616
Farm income	1.581	2.304
Farm has livestock	-13.115	5.747
Farm has crop insurance	3.089	6.006
Field Characteristics		
Field acres	-0.006	0.041
Field is rented	-0.234	5.596
Soil texture is clay	4.294	6.329
Soil texture is sand	6.145	19.520
Distance to Lake Erie	0.000	0.000
Slope is greater than 2%	19.768	29.637
Field has good soil	71.249**	37.865
Field has poor soil	-15.589	37.650
Dummies for crop rotation and P frequency choices		
Dummy - corn multi	12.288*	6.545
Dummy - soybean single	2.462	7.774
Dummy - soybean multi	20.268**	9.582
Dummy - other crop	-4.026	11.600
Intercept	233.527***	40.900
Implied mean elasticity	*-1.568****	

Table 28. Descriptive Evidence on Heterogeneity in Phosphorus Price Responsiveness - Ordinary Least Squares Regression

	25th quantile	50th quantile	75th quantile
Input and Output Prices			
Normalized P fertilizer price	-0.3423**	-0.7550***	-1.0397***
	(0.1499)	(0.2768)	(0.3801)
Normalized N price	0.2253***	0.4644***	0.5275***
	(0.0571)	(0.1054)	(0.1448)
Normalized expected corn price	0.0408	0.0872	-0.3080
	(0.1533)	(0.2832)	(0.3888)
Normalized manure price	-1797.74	-5489.62	-5155.17
	(2440.14)	(4506.32)	(6187.33)
Farmer Characteristics			
Familiar with 4R Nutrient	-0.0477	-1.7467	-3.9712
stewardship	(1.3614)	(2.5142)	(3.452)
More risk loving	-0.3524	-3.3623**	-3.762**
	(0.7237)	(1.3365)	(1.835)
Age	-0.0021	-0.1910	-1.0943**
	(0.1794)	(0.3313)	(0.4549)
Education	0.3298	1.3635	-4.3664
	(1.2559)	(2.3193)	(3.1844)
Years of farming experience	0.0702	0.2188	0.3654
	(0.1155)	(0.2133)	(0.2929)
Female operator	-2.3411	-81.371	-81.371
	(20.851)	(38.506)	(52.870)
Farm income	0.1523	-2.3012	-2.454
	(1.4729)	(2.7200)	(3.735)
Farm has livestock	-2.3458	-20.916***	-11.038
	(3.6630)	(6.765)	(9.288)
Farm has crop insurance	0.1636	2.8488	4.899
	(3.6630)	(7.0805)	(9.721)
Field Characteristics			
Field acres	-0.0077	-0.0098	0.0065
	(0.0259)	(0.0478)	(0.0656)
Field is rented	2.1760	1.5745	-2.9959
	(3.5704)	(6.5936)	(9.0532)

Continued

Table 29. Descriptive Evidence on Heterogeneity in Phosphorus Price Responsiveness - Quantile Regressions

Table 29 continued

	25th quantile	50th quantile	75th quantile
Field Characteristics			
Soil texture is clay	0.6861	6.698	3.574
	(4.0415)	(7.4637)	(10.248)
Soil texture is sand	8.8723	14.732	18.882
	(12.4688)	(23.027)	(31.617)
Distance (meter) to Lake Erie	-1.57E-05	-0.0001	-0.0001
	(4.01E-05)	(0.0001)	(0.0001)
Slope is greater than 2%	-0.4608	-2.2668	-4.8444
	(3.4496)	(6.3706)	(8.747)
Field has good soil	-0.0828	6.6880	14.826
	(4.0146)	(7.4139)	(10.180)
Field has poor soil	0.1894	10.3713	11.675
	(4.1609)	(7.6842)	(10.551)
Crop and P Fertilizer Frequency Choices			
Dummy – corn multi	1.7705	10.925	6.820
	(4.1833)	(7.7256)	(10.607)
Dummy – soybean single	-0.4259	-5.2671	0.6612
	(4.9458)	(9.1336)	(12.541)
Dummy – soybean multi	-1.7932	22.7628**	52.669***
	(6.1043)	(11.2731)	(15.478)
Dummy – other crop	-1.4384	-7.7552	-18.578
	(7.3940)	(13.655)	(18.749)
Intercept	18.774	99.010***	272.237***
	(17.234)	(31.827)	(43.700)
Number of observations		1136	
Implied mean elasticity	-3.0030**	-0.8477***	-0.5364***
Implied average P reduction (lbs/ac) for a 10% price increase	-0.1966**	-4.3366***	-5.972***
	(0.0861)	(1.5900)	(2.183)

	corn single	corn multi	soybean single	soybean multi
Yield Equation				
Input and Output Prices				
Normalized N price	0.0411	-0.063**	0.0436	0.0414
	(0.0327)	(0.0327)	(0.0379)	(0.0296)
Normalized P price	0.306**	-0.084	0.1288	-0.0341
	(0.1756)	(0.146)	(0.1092)	(0.1563)
Normalized expected corn	-0.1499	-0.1968	-0.045	0.0222
price	(0.1693)	(0.1663)	(0.1025)	(0.1068)
Normalized manure price	2932.62	371.19	28056***	1272.1
	(3021.4)	(1894.3)	(7146.6)	(3546.8)
Farmer Characteristics				
Familiar with 4R	0.6166	1.981	-2.208**	0.6527
	(1.622)	(1.407)	(0.932)	(1.026)
More risk loving	1.9505	2.1753***	-0.219	0.1240
	(4.3008)	(0.8319)	(0.438)	(0.5366)
Age	-0.268	-0.291	0.0312	-0.0263
	(0.3591)	(0.248)	(0.1525)	(0.1721)
Education	1.5744	-3.173**	1.0346	2.796***
	(1.6413)	(1.612)	(0.8138)	(0.9967)
Years of farming experience	0.0123	-0.073	-0.0779	0.1170
	(0.1082)	(0.217)	(0.1253)	(0.1711)
Female operator	-41.531	-9.177	0.2241	1.3402
	(29.445)	(27.725)	(12.955)	(7.9760)
Farm income	3.0790*	0.833	0.5852	0.1768
	(1.7396)	(1.704)	(0.9221)	(1.0303)
Farm has livestock	-0.9386	-3.167	3.784*	2.7275
	(4.1414)	(3.755)	(2.295)	(2.9921)
Farm has crop insurance	1.5569	-2.377	1.891	-3.4003
	(4.5892)	(4.229)	(2.192)	(2.6616)
Field Characteristics				
Field acres	-0.0159	0.0168	0.0489	0.0461
	(0.0262)	(0.0265)	(0.0382)	(0.0358)

Continued

Table 30. SUREG Regression Results for Yield, Nitrogen and Manure Equations with Bootstrapped Standard Errors for Table 19

Table 30 continued

	corn single	corn multi	soybean single	soybean multi
Field is rented	1.9505	-6.737*	0.2030	-2.2713
	(4.3008)	(3.769)	(2.2522)	(2.5506)
Soil texture is clay	-5.9174	-11.479*	-2.792	-3.9099
	(5.9255)	(3.769)	(2.134)	(2.5492)
Soil texture is sand	0.5592	-15.110	1.1142	-7.6139
	(16.026)	(10.902)	(7.4025)	(10.919)
Distance (meter) to Lake	-0.0001**	-4.4E-05	-5.2E-05**	7.35E-05**
Erie	(4.6E-05)	(4.6E-05)	(2.4E-05)	(2.93E-05)
Slope is great than 2%	2.2490	-7.336**	-0.3473	-3.3196
	(4.7858)	(3.609)	(2.0355)	(2.3380)
Field has top soil	11.295**	6.556*	5.525**	-4.185
	(4.7858)	(3.977)	(2.630)	(2.981)
Field has poor soil	-3.936	-2.096	-3.850*	2.359
	(5.134)	(4.837)	(2.290)	(2.935)
Inverse Mills ratio for crop and P frequency	-10.054	-21.29***	-0.430	4.052
choices	(10.541)	(8.260)	(2.346)	(2.946)
Intercept	148.30***	231.18***	45.336***	27.994*
	(19.648)	(26.890)	(11.178)	(15.151)
Nitrogen Fertilizer Demand Equation				
Input and Output Prices				
Normalized N price	0.2896***	-1.245**	0.6469***	0.995***
	(0.0835)	(0.581)	(0.1415)	(0.2922)
Normalized P price	-0.526	0.258**	0.1753	0.4319
	(0.4470)	(0.130)	(0.4076)	(1.5000)
Normalized expected corn	0.0704	-0.471	0.5071	0.6943
price	(0.4320)	(0.6763)	(0.3829)	(1.0556)
Normalized manure price	-3246.44	7238.88	88752***	3568.6
	(7708.5)	(7699.7)	(26691)	(35063)
Farmer Characteristics				
Familiar with 4R	-4.7008	-4.985	2.926	1.883
	(4.1383)	(5.721)	(3.483)	(10.141)
More risk loving	0.0244	-9.239***	-1.300	2.766
	(2.2364)	(3.382)	(1.636)	(5.305)

Continued

Table 30 continued

	corn single	corn multi	soybean single	soybean multi
Age	0.1360	-0.796	-0.1734	-2.106
	(0.5075)	(1.010)	(0.5697)	(1.702)
Education	1.7269	-6.654	-6.296**	-15.561
	(4.1874)	(6.554)	(3.040)	(9.853)
Years of farming experience	-0.1596	-0.842	0.1372	1.988
	(0.2579)	(0.882)	(0.4682)	(1.692)
Female operator	-58.963	-163.55	-23.106	-37.147
	(75.125)	(112.73)	(48.394)	(78.849)
Farm income	9.6162*	8.9147	-1.541	3.3305
	(4.4383)	(6.927)	(3.444)	(10.186)
Farm has livestock	-0.8520	-33.60**	-9.322	-24.663
	(10.566)	(15.267)	(8.573)	(29.579)
Farm has crop insurance	9.5273	2.7665	9.175	36.644
	(11.708)	(17.194)	(8.186)	(26.312)
Field Characteristics				
Field acres	-0.0759	-0.176	0.1162	-0.402
	(0.0668)	(0.1075)	(0.1427)	(0.3541)
Field is rented	-8.8139	7.6407	-9.598	6.8117
	(10.972)	(15.324)	(8.413)	(25.215)
Soil texture is clay	16.2418	-21.384	-8.172	-16.746
	(15.1180)	(17.508)	(7.970)	(25.201)
Soil texture is sand	-13.6574	-81.674*	65.291**	-22.541
	(40.889)	(44.328)	(27.652)	(107.94)
Distance to Lake Erie	9.53E-05	-0.0002	-3.7E-05	1.42E-04
	(0.0001)	(0.0002)	(9.0E-05)	(0.0003)
Slope is great than 2%	-27.360***	-30.381**	0.7487	54.480**
	(10.5202)	(14.673)	(7.6035)	(23.113)
Field has top soil	3.7291	10.066	2.3065	65.398**
	(12.2100)	(16.170)	(9.8228)	(29.466)
Field has poor soil	15.6949	-3.185	6.7582	-43.076
	(13.097)	(19.665)	(8.5533)	(29.011)
Inverse Mills ratio for	-74.530	-44.427	2.1581	-11.763
crop and P frequency choices	(10.541)	(33.583)	(8.5533)	(29.127)
Intercept	193.89***	460.21***	6.3878	31.806
	(50.112)	(109.13)	(41.743)	(149.79)

Continued

Table 30 continued

	corn single	corn multi	soybean single	soybean multi
Manure Demand Equation				
Input and Output Prices				
Normalized N price	-2.0205	-3.485	0.3559	-1.3033
	(7.5040)	(7.307)	(2.5340)	(1.7215)
Normalized P price	-179.23***	-41.124	1.2916	-3.5422
	(40.201)	(33.284)	(7.2993)	(9.3136)
Normalized expected corn	21.985	2.315	-38.791***	0.5106
price	(38.552)	(37.824)	(6.873)	(6.2209)
Normalized manure price	-2.85E07**	0.55E06		
	(1.3E07)	(0.8E06)		
Dummy for zero normalized	-14347***	-1327	-5505.7***	-5448.5***
manure price	(4988)	(2767.8)	(956.35)	(883.11)
Number of dairy cows	3.713	5.983	4.306*	0.0008
	(3.534)	(16.602)	(2.339)	(0.0067)
Number of poultry	-0.0007	-0.0095	-0.2187	0.0175
	(0.0327)	(0.0393)	(0.4341)	(0.0332)
Farmer Characteristics				
Familiar with 4R Nutrient	184.53	597.16*	49.199	34.129
Stewardship	(369.60)	(325.72)	(62.713)	(60.606)
More risk loving	204.41	136.38	-5.699	-63.17**
	(199.70)	(190.39)	(29.606)	(31.40)
Age	-110.73**	-2.818	-6.802	-17.026*
	(45.33)	(56.538)	(10.206)	(10.088)
Education	-69.79	-479.24	-29.999	-12.665
	(374.19)	(369.52)	(54.735)	(58.636)
Years of farming experience	-6.155	-9.328	4.528	15.915
	(24.631)	(49.248)	(8.387)	(10.037)
Female operator	1765.98	1213.56	-217.24	-297.63
	(6702.9)	(6298.4)	(865.58)	(464.86)
Farm income	-511.26	15.140	-36.996	-149.83**
	(403.42)	(391.99)	(61.668)	(61.445)
Farm has livestock	1831.8*	-281.16	348.07**	-211.12
	(966.63)	(876.81)	(160.06)	(185.30)
Farm has crop insurance	-1873.7*	706.82	-121.12	-12.95
	(1047.04)	(960.81)	(147.37)	(155.07)

Continued

Table 30 continued

	corn single	corn multi	soybean single	soybean multi
Field Characteristics				
Field acres	3.1480	-7.306	3.459	0.3278
	(5.9777)	(6.346)	(2.555)	(2.0917)
Field is rented	317.06	-973.42	-269.64*	-74.312
	(985.82)	(859.52)	(151.02)	(150.07)
Soil texture is clay	375.20	816.71	-175.03	-143.58
	(1357.1)	(986.80)	(143.10)	(149.76)
Soil texture is sand	-1780.2	-1056.3	-48.797	212.42
	(3648.8)	(2494.9)	(494.44)	(635.98)
Distance to Lake Erie	0.0221**	0.0045	0.0023	0.0040
	(0.0106)	(0.0107)	(0.0016)	(0.0017)
Slope is great than 2%	255.81	-793.81	-62.015	-78.097
	(940.96)	(829.52)	(136.19)	(139.22)
Field has top soil	4.757	1713.55*	154.52	-343.79*
	(1097.52)	(921.8)	(175.64)	(177.13)
Field has poor soil	640.16	440.90	330.11**	-134.63
	(1170.22)	(1108.6)	(154.28)	(174.24)
Inverse Mills ratio for	3015.7	-1412.23	-230.11	-63.32
crop and P frequency choices	(2417.43)	(1900.9)	(157.24)	(172.10)
Intercept	27112***	5653.10	7796.1***	6939.5***
	(6934.56)	(6703.65)	(1094.6)	(928.63)

	corn single	corn multi	soybean single	soybean multi
Phosphorus Fertilizer Prices with Interactions - Targeting				
Normalized P fertilizer price	-0.4312	-3.718***	-2.4962	-2.1109
	(0.7096)	(1.1549)	(1.8569)	(4.5265)
Normalized P price * familiar 4R	0.4642	1.3910***	-0.0847	0.3661
	(0.2829)	(0.3973)	(0.6079)	(1.8682)
Normalized P price * slope	-1.139*	-1.6506	1.4634	-8.4639
	(0.6472)	(1.4531)	(1.7230)	(7.2907)
Normalized P price * good soil	0.3317	-1.9769*	-0.889	5.9426
	(0.8107)	(1.0998)	(2.0961)	(3.6671)
Normalized P price * poor soil	-1.0315	1.5477	0.6032	2.4985
	(0.8195)	(1.6736)	(1.7805)	(3.5544)
Input and Output Prices				
Normalized N price	0.3241***	0.1997	0.4641	-0.2932
	(0.4220)	(0.2113)	(0.4092)	(0.4127)
Normalized expected corn price	0.4220	-1.2744**	0.2352	-2.105***
	(0.3086)	(0.5695)	(0.5807)	(0.7469)
Normalized manure price	-3248.27	6394.64	-26438	23278
	(5446.32)	(6339.23)	(63994)	(27763)
Farmer Characteristics				
Familiar with 4R Nutrient stewardship	-29.403*	-87.800***	13.136	-33.067
	(16.764)	(23.508)	(35.341)	(111.60)
More risk loving	-3.653**	-9.542***	-1.423	5.1667
	(1.573)	(2.800)	(2.519)	(3.9432)
Age	-0.258	-0.9790	-0.1210	1.6303
	(0.3588)	(0.8400)	(0.8276)	(1.2292)
Education	-0.5988	-6.6612	-3.4279	-24.566***
	(2.9129)	(3.5036)	(4.6296)	(7.024)
Years of farming experience	-6.735**	0.2211	-0.4468	-2.5880**
	(0.1938)	(0.7270)	(0.7295)	(1.2226)
Female operator	-74.058	-55.543	454.69***	13.8162
	(52.734)	(92.565)	(74.648)	(55.749)
Farm income	-6.735**	6.555	3.9731	14.169*
	(3.1427)	(5.6948)	(5.3122)	(7.5300)

Continued

Table 31. SUREG Regressions for Phosphorus Fertilizer Demand without Constraining the Mean Elasticity Coefficient from Reduced-form Panel Data Model

Table 31 continued

	Corn single	Corn-multi	Soybean-single	Soybean-multi
Farmer Characteristics				
Farm has livestock	-14.161*	-11.420	-17.386	11.129
	(7.490)	(12.611)	(13.005)	(21.512)
Farm has crop insurance	2.580	14.262	-14.083	42.714**
	(8.2340)	(14.207)	(12.403)	(18.825)
Field Characteristics				
Field acres	-0.0282	-0.1184	0.0567	-0.5524**
	(0.0471)	(0.1012)	(0.2174)	(0.2522)
Field is rented	2.1116	-1.9683	-6.1786	-1.5923
	(7.7066)	(12.605)	(12.881)	(17.911)
Soil texture is clay	19.819*	-10.656	3.2072	2.013
	(10.690)	(14.266)	(12.246)	(17.8886)
Soil texture is sand	36.493	-31.394	22.136	-158.41**
	(28.720)	(36.972)	(41.762)	(76.581)
Distance (meter) to Lake Erie	-0.0001	-0.0004***	9.12E-05	-3.8E-04*
	(0.0000)	(0.0002)	(0.0001)	(0.0002)
Slope is greater than 2%	74.120**	97.728	-96.232	493.05
	(37.627)	(86.459)	(101.65)	(436.96)
Field has good soil	-2.145	116.48*	69.118	-372.22*
	(47.485)	(64.819)	(124.84)	(217.84)
Field has poor soil	72.99	-79.142	-29.829	-178.83
	(47.643)	(99.581)	(104.53)	(214.07)
Inverse Mills ratio for 1st stage	-22.182	-56.555**	18.761	5.523
crop and P frequency choices	(19.046)	(28.303)	(14.222)	(21.479)
Intercept	134.06***	575.12***	217.01*	385.96***
	(49.518)	(108.66)	(119.65)	(278.51)
Implied mean elasticity	-0.3548	-2.6200***	-2.2002	-1.492

	corn single	corn multi	soybean single	soybean multi
Constrained Phosphorus Fertilizer Prices from Reduced-form Regressions				
Normalized P fertilizer price	-0.4376*	-0.5634***	-0.4104***	-0.8462***
	(0.2259)	(0.1689)	(0.1111)	(0.2325)
Phosphorus Fertilizer Prices with Interactions - Targeting				
Normalized P price * familiar 4R	0.4685*	0.7552***	-0.4593	-0.0464
	(0.2593)	(0.3274)	(0.4784)	(1.1034)
Normalized P price * slope	-0.8472	-2.1879	1.892*	-8.9037
	(0.6308)	(1.4607)	(1.1254)	(7.1390)
Normalized P price * good soil	0.0369	-3.906***	-2.020	5.0280*
	(0.6992)	(0.8573)	(1.9778)	(2.5980)
Normalized P price * poor soil	-0.7985	-0.4070	-1.0069	0.7629
	(0.6616)	(1.5486)	(0.7915)	(1.4658)
Input and Output Prices				
Normalized N price	0.3117***	0.1889	0.5170	-0.3087
	(0.1196)	(0.2152)	(0.4094)	(0.4055)
Normalized expected corn price	0.3960	-1.0013*	0.2344	-2.1253***
	(0.3059)	(0.5724)	(0.5808)	(0.7485)
Farmer Characteristics				
Familiar with 4R Nutrient	-29.588*	-50.931***	18.136	-8.1610
stewardship	(15.463)	(19.721)	(27.748)	(64.913)
More risk loving	-3.757**	-9.542***	-1.7647	5.6304
	(1.552)	(2.844)	(2.5021)	(3.8592)
Age	-0.304	-1.1262	-0.2722	1.5946
	(0.3585)	(0.8539)	(0.8669)	(1.2031)
Education	-0.5537	-7.5660	-3.4345	-25.609***
	(2.9109)	(5.3865)	(4.6233)	(6.941)
Years of farming experience	0.2910	0.1859	-0.2509	-2.6831**
	(0.1937)	(0.7402)	(0.7168)	(1.1908)
Female operator	-75.761	-68.220	462.04***	16.6523
	(52.711)	(94.177)	(73.742)	(55.823)
Farm income	-6.732**	5.170	4.2035	14.423*
	(3.1409)	(5.778)	(5.3208)	(7.485)

Continued

Table 32. SUREG Regression Results for Phosphorus Fertilizer Demand Without Including Manure Demand and Manure Prices

Table 32 continued

	Corn single	Corn-multi	Soybean-single	Soybean-multi
Farmer Characteristics				
Farm has livestock	-14.823**	-12.991	-16.626	11.368
	(7.406)	(12.779)	(13.031)	(21.563)
Farm has crop insurance	2.235	14.072	-14.051	45.757**
	(8.2206)	(14.467)	(12.450)	(18.556)
Field Characteristics				
Field acres	-0.0320	-0.0590	0.0539	-0.5597**
	(0.0471)	(0.0963)	(0.2165)	(0.2516)
Field is rented	2.2286	-2.6392	-4.1377	-4.0456
	(7.6926)	(12.761)	(12.838)	(17.748)
Soil texture is clay	20.220*	-13.300	1.0129	0.0639
	(10.685)	(14.457)	(12.202)	(17.6339)
Soil texture is sand	35.666	-28.121	26.335	-162.98**
	(28.707)	(37.493)	(41.797)	(76.51)
Distance (meter) to Lake Erie	-0.0001	-0.0004***	7.9E-05	-3.6E-04*
	(0.0001)	(0.0002)	(0.0001)	(0.0002)
Slope is greater than 2%	57.385	126.83	-121.38	517.35
	(36.684)	(87.12)	(66.04)	(427.75)
Field has good soil	14.843	230.04*	136.23	-318.09**
	(41.507)	(50.654)	(117.98)	(152.61)
Field has poor soil	60.675	36.167	65.074	-71.913
	(38.504)	(92.387)	(46.178)	(87.513)
Inverse Mills ratio for 1st stage	-23.982	-57.783**	16.962	6.938
crop and P frequency choices	(18.997)	(28.802)	(14.228)	(21.322)
Intercept	141.14***	397.65***	100.64*	316.57***
	(32.738)	(89.444)	(53.016)	(84.058)
Implied mean elasticity	-0.3600***	-0.3970***	-0.3617***	-0.5982***

State	Farm Acre Groups	2007 Census of Agriculture Farms	Acres	% acres	Our Farmer Survey Farms	Acres	% acres
Indiana	< 10 acres	316	1450	0.26%			
Indiana	10-49 acres	1269	21383	3.82%	16	547	0.29%
Indiana	50-179 acres	1104	80555	14.38%	80	9259	4.98%
Indiana	180-499 acres	501	133386	23.80%	101	32710	17.59%
Indiana	500-999 acres	165	107323	19.15%	60	39474	21.22%
Indiana	1000-1999 acres	94	127018	22.67%	35	47597	25.59%
Indiana	> 2000 acres	34	89253	15.93%	22	56397	30.32%
Indiana Subtotal		3483	560368	100%	314	185984	100.00%
Michigan	< 10 acres	40	120	0.02%	1	8	0.01%
Michigan	10-49 acres	689	10403	2.01%	3	100	0.13%
Michigan	50-179 acres	849	49423	9.55%	16	1721	2.23%
Michigan	180-499 acres	393	94790	18.32%	17	5527	7.18%
Michigan	500-999 acres	166	101770	19.67%	24	17111	22.22%
Michigan	1000-1999 acres	112	136784	26.44%	20	25958	33.71%
Michigan	> 2000 acres	47	124037	23.98%	8	26585	34.52%
Michigan Subtotal		2296	517327	100%	89	77010	100%
Ohio	< 10 acres	251	1005	0.03%	1	8	0.00%
Ohio	10-49 acres	2127	42659	1.18%	74	2263	0.25%
Ohio	50-179 acres	3178	271555	7.50%	306	33804	3.76%
Ohio	180-499 acres	2230	620761	17.15%	344	106243	11.82%
Ohio	500-999 acres	997	645401	17.83%	215	149786	16.67%
Ohio	1000-1999 acres	528	676477	18.69%	128	172838	19.24%
Ohio	> 2000 acres	103	284746	7.87%	59	170587	18.99%
Ohio Subtotal		9414	2542604	100%	1127	635530	100%
Three States' Total		15193	3620298		1530	898524	

Table 33. Comparison of Farm Acre Distribution between Our Farmer Survey and 2007 Census of Agriculture Microdata

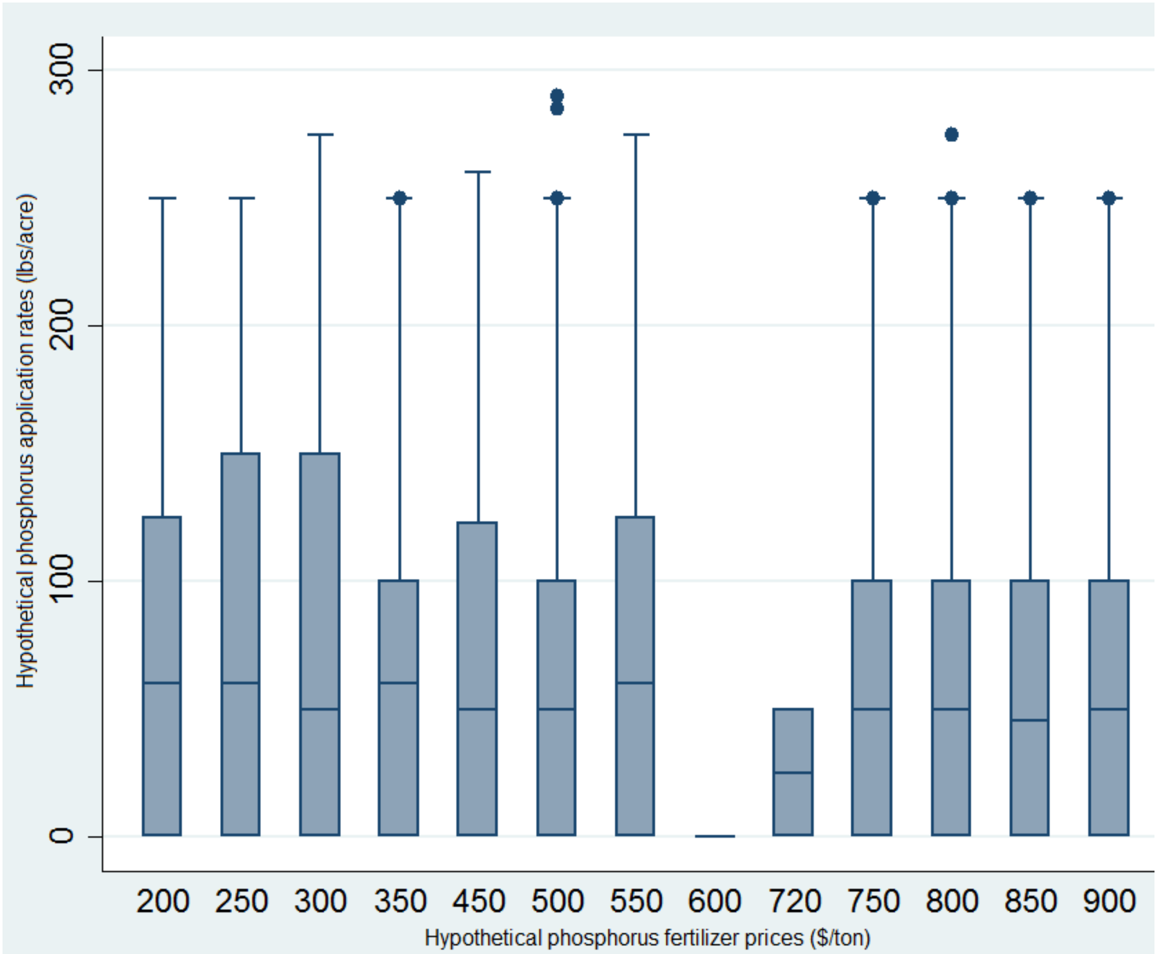


Figure 13. Distribution of Fertilizer Application Rates Based on Responses to Hypothetical Fertilizer Price Questions