Kevin MEYER*

The Impact of Agricultural Land Use Change on Lake Water Quality: Evidence from Iowa

The environmental impacts of agricultural policies must be quantified to perform full cost-benefit analyses and make informed policy decisions. In this paper I use a unique panel data set to estimate the effect of changes in cropland on lake water quality. Fifteen years of water quality measurements across over 100 lakes are combined with satellite imagery and weather data. Using a dynamic panel data model, I find that the elasticity of water quality to cropland is 0.0535. To understand the policy implications, I estimate a second model to find the elasticity of cropland to crop prices. I combine these estimates to analyze the effect of the Renewable Fuel Standard (RFS). I find that the RFS decreased lake water quality; however, the magnitude of this effect is negligible.[1]

Keywords: Agriculture, land use, water quality

JEL classifications: Q15, Q18

* Assistant Professor, Department Of Economics, Saginaw Valley State University, 321 Curtiss Hall, Saginaw MI, 48642 kmeyer@svsu.edu

Received: 25 January 2018; Revised: 14 May 2018; Accepted: 23 May 2018

Introduction

Understanding the effect of land use on water quality is an important question in environmental policy. Agriculture is consistently identified by the EPA as a cause of water quality degradation due to fertilizer runoff entering nearby waterways (EPA, 2002). These effects have important economic and ecological consequences, such as adverse effects on recreation and drinking water quality. As has been known since the time of Pigou, these effects are a classic case of a market failure (Pigou, 1920). Since water quality is not a market good, the free market will not result in an optimal level of lake water quality. Efficient government intervention requires understanding the benefits and costs of land use on water quality, however measuring these effects has been difficult, primarily due to a lack of suitable data.

The primary research objective of this paper is to estimate the elasticity of lake water quality to land use change. These estimates can be used by future researchers and policymakers as part of cost benefit analyses, where policies affect land use change near lakes. My approach combines high quality water measurements across 100 Iowa Lakes over 15 years, along with satellite data on cropland use and PRISM data on weather. As an extension, I estimated a secondary model of the elasticity of land use change to crop prices. I use the estimates from these two models to estimate the effect of the Renewable Fuel Standard (RFS) on lake water quality. Due to the inelastic estimates from both models, I find a negligibly positive impact of the RFS on lake water quality.

This paper contributes to the literature in several ways. First, it exploits 15 years of water quality measurements across Iowa to perform a statistical analysis of the effect of land use on water quality. Second, it provides strong statistical evidence of a persistent effect of water quality across time. Third, it adds to the literature on the response of cropland expansion to crop prices. Finally, it adds to the literature on the environmental effects of biofuel related policies.

The paper is organized as follows. The next section provides some background on the typical techniques used to assess the relationship between land use and water quality. This is followed by a description of the econometric model used in both the water quality and cropland response models. I then give a detailed description of the data set, followed by the results for both models. These results are then used in an application to estimate the effect of the RFS on lake water quality, followed by a summary.

Background

There is a long history of studies that attempt to identify the relationship between land use and the water quality of lakes, rivers, and streams. Most of these studies can be divided into two types- simulation models such as SWAT and BASINS[2], and econometric models. The former are able to model complex relationships between the climate, land use, and water quality to examine issues that might otherwise be intractable. For example, simulations from these types of programs have been used to examine the hypoxia “dead zone” in the Gulf of Mexico (Rabotyagov, 2014), the effect of corn-based ethanol on environmental quality (Secchi, 2009), and the potential for cropland to reduce flood risk (Schilling, 2014). Simulation models are invaluable for gaining insight into issues that may otherwise be too complicated for any one statistical model to capture, but they have drawbacks. On a practical level, the complexity of the simulated relationships requires many parameters, and choosing these parameters requires a significant amount of expertise. This can make it difficult for other researchers to truly understand what is generating the results. Statistical models, on the other hand, are helpful in their ability to model relationships between variables in a relatively straightforward and transparent way.

Many statistical analyses in the literature rely on simple correlation coefficients between different land uses and a

[1] I would like to thank the EPA for their support of related work as a graduate student at Iowa State University. In addition, thanks to David Keiser, Catherine L. Kling, Gian Carlo Moschini, Quinn Weninger, and John Beghin for their helpful comments.

[2] SWAT stand for Soil and Water Assessment Tool; BASINS stands for Better Assessment Science Integrating point & Non-point Sources.

measure of water quality. For example, Tong (2002) found a statistically significant positive correlation of 0.1913 and 0.1563 between agriculture and total nitrogen and phosphorous in surface water, respectively, in 11-digit HUCS in Ohio. In fact, most studies find a positive correlation between the two variables (Meador, 2003; Dauer, 2000). While correlations are informative, they do not help isolate causal effects. In other words, the question arises: does an increase in cropland cause the water quality to drop, or could it represent something else, such as the quality of the land? Answering this requires a model that controls for the quality of land, as well as other possible omitted variables.

Many studies have used regression techniques to try to estimate the relationship. Tu (2011) uses geographically weighted regressions to estimate local effects in an area surrounding Boston. He estimates a separate univariate regression for 6 different water quality variables and 14 land uses, for a total of 84 regressions. The results showed little influence of agricultural land on water quality. A drawback of this study is that water quality measurements are averaged over time and estimated using only one year of observed land uses. In fact, cross-sectional regressions are common in water quality studies – possibly due to a lack of quality, publicly available time series data. Another technique used in the literature is to use simple univariate regressions of land use on water quality (see, e.g. Lougheed, 2001). Limiting the model to one period, or not controlling for other factors that can affect water quality can potentially bias the coefficients of interest.

This omitted variable bias problem can potentially create misleading results. For example, Sprague (2012) studies the effect of the Conservation Reserve Program (CRP) on total nitrogen and phosphorus loadings in rivers using a cross-sectional regression and found a marginally positive effect, indicating that CRP land increases nutrient levels; this is the opposite effect intended by the program and lacks a credible explanation. The key problem in this study is that, for the results to hold, the model must assume that CRP land is randomly distributed across space and uncorrelated with omitted factors that affect water quality. This is unlikely to be the case since profit maximizing landowners will choose to retire the least profitable farmland into the CRP first. In Iowa, for example, CRP land is concentrated in the south, where the soil quality is relatively low. Therefore, it would not be surprising to find a negative correlation between CRP land and water quality, since lower quality soil typically means increased runoff.

The geographic characteristics of water bodies has led to some studies that use more complicated regression models. Atasoy (2006) employs spatial econometric techniques to study the effect of urban land use on water quality. Their analysis uses monthly nutrient measurements over a four-year period combined with monthly measures of urban development, weather variables, and a single year of satellite imagery to control for agricultural land. Their emphasis on rivers and streams is an example of how geography plays an important role in the specification of an appropriate econometric model when studying environmental issues. In their study, upstream river quality clearly affects downstream river quality as it is carried through a stream network, thus it makes sense to explicitly include a spatially-lagged dependent variable while allowing for temporal correlations in the error term. In this study, where the observed unit is lakes, it does not make sense conceptually to include a spatial lag, since lakes do not flow into each other. Instead, it is appropriate to include a temporal lag of the dependent variable, since lake water remains relatively stationary over time. This implies that nutrient levels may persist; this effect is known as the "hydraulic retention time". As an example, Jeppesen (2005) reduced the nutrient levels in multiple lakes and observed that the lakes did not reach a new, lower steady state for 15 years.

Existing lake water quality studies that attempt to include dynamics have typically been confined to one lake and its watershed. For example, Balkcom (2003) use multiple samples from a lake over time to calibrate an integrated assessment model, which was then used to analyse different land use scenarios. By contrast, this study uses data on over 120 lakes over 15 years, creating a rich variation in lake quality, geographical characteristics, and the characteristics of surrounding land use.

As one of the most productive farming states in the country, Iowa land use can be particularly sensitive to changes in farming policies. Therefore, given the evidence of the link between cropland and water quality degradation, government policies can directly and indirectly affect water quality. Two primary examples are the Conservation Reserve Program (CRP), and the Renewable Fuel Standard (RFS). The CRP has evolved from its initial goals of removing cropland from production to focus more on maximizing the environmental impact of the program. Only land currently in production or expiring CRP land are eligible to be retired and receive CRP subsidies, and retired land must be planted with species that will improve environmental health and quality. Thus, the possible water quality benefits of an acre of CRP land are 1) removing a hectare of cropland, and therefore all related nutrient use, and 2) replacing it with a hectare plants that can help improve soil quality and reduce runoff of nutrients from the surrounding area. CRP land in Iowa began a major decline around 2007. It is likely that multiple factors contributed to this decline, especially rising crop prices (and thus profitability of land) and a decline in funding for the program.

The RFS, first established under the Energy Policy Act of 2005, mandated 28 billion liters of ethanol be used by 2012. The scope of the biofuel mandate expanded significantly in 2007 by mandating 136 billion liters of ethanol in the U.S. by 2022. Most of the current biofuel supply comes from corn ethanol. Therefore, the biofuel mandate has and will continue to have significant economic and environmental impacts on Iowa, the nation's leading corn producer. Researchers have identified water quality degradation as an important consequence of biofuel production (Simpson, 2008). Although corn cultivation requires a significant amount of water, water shortages are typically not a concern in Iowa. Rather, the increased use of nutrients from expanding corn production along both the intensive and extensive margin are of concern. In addition, an increase in the demand for corn can affect the price of other crops, such as soybeans, which can cause cropland expansion for those crops as well. As corn uses nitrogen relatively inefficiently (Balkcom, 2003), switching over to

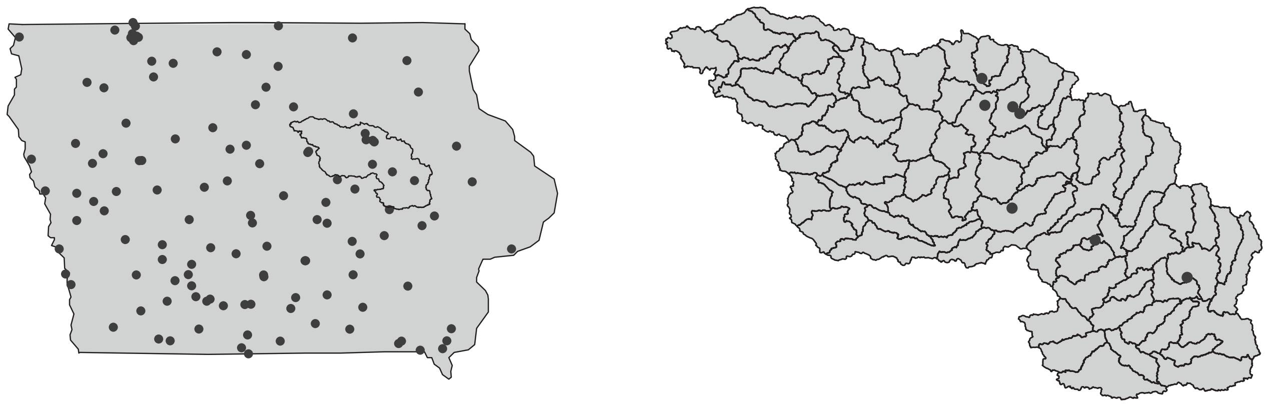

Figure 1: Location of Iowa lakes, represented by solid dots (left), a HUC 8 watershed comprised of smaller HUC12 watersheds (right).

Source: Iowa State Limnology Department, retrieved from http://limnology.eeob.iastate.edu/lakereport/default.aspx

corn from other crops can potentially increase the amount of nitrogen in the soil. Finally, if we assume that farmers grow crops on the best farmland available, cropland expansion will likely occur in marginal, more environmentally sensitive areas, including CRP land (Secchi, 2009). Thus, the two policies mentioned here are to some degree interdependent, as farmer's will look to maximize their profit by either accepting subsidies to retire the land into the CRP, or to farm the land and sell their crops.

Data

Water quality data were downloaded from Iowa State's Limnology Laboratory website [3]. The Iowa Lakes data uses consistent, scientifically based, and well documented hydrological sampling methods. In this paper I use the average of three annual measurements. Averaging three lake water samples over a year offers adequate precision for water quality indicators (Downing, 2006). CTSI, a measure of lake water clarity, is used as the main water quality indicator because it summarizes the outcome of increased sediment or nutrient loadings, as opposed to a measurement of the inputs of sediments or nutrients into a lake. Most lakes have a CTSI between 0-100, with each increase of 10 units representing an approximate doubling in algal biomass. An intuitive way to think about the index is that a CTSI of 0 represents a visible depth of 64 meters, while a CTSI of 100 represents a visible depth of only 6.4 centimeters. The CTSI can be approximately divided into four trophic classes: oligotrophic (less than 30-40), mesotrophic (40-50), eutrophic (50-70), and hypereutrophic (70-100+). The left panel of Figure 1 shows the locations of the 123 lakes used in this study.

Annual land use data comes from USDA NASS cropland data layers (CDL's), which are satellite images. Each pixel of a satellite image is assigned a land use based on color analysis. I find the total land use for a geographic region by summing the pixels assigned to each land use. Since the focus of the paper is water quality, land use was aggregated to the local watershed level, known as a HUC (hydrologic unit code). Aggregating to a watershed captures drainage characteristics more accurately than aggregating to an arbitrary governmental boundary. HUC's differ in size and are nested within each other- a HUC12 is located within a HUC10, which itself is within a HUC8, and so on. The right panel of Figure 1 shows the size of typical HUC12 watersheds, as well as the larger HUC8 which contains them.

For this paper I focus on land devoted to corn, soybeans, and grassland[4]. An issue with using cropland devoted to corn and soybean use is they are strongly positively correlated, with a correlation coefficient of .80. To avoid multicollinearity, I sum these two land uses into a single variable labeled *crops*. Since official CRP enrollment numbers are only available at the county level, I include grassland as a proxy for the effect of CRP land on water quality.

Data for precipitation and temperature were calculated from Oregon State's PRISM dataset[5]. PRISM provides the daily precipitation and temperature for 30km by 30km grid cells that cover the continental U.S. To find the annual precipitation for an individual HUC12 I sum the daily data for each PRISM grid cell across the watershed, and then sum the daily values over the year. To find the average annual temperature for each HUC12 I average daily temperatures across PRISM grid cells, and then average the daily values over the year.

Table 1 provides a description and summary statistics for the variables included in the analysis.

Empirical Models

Water Quality Model

I use the following dynamic panel data model to estimate the effects of land use on water quality:

$$Q_{\{i,j,t\}} = \mu_i + B_1 Q_{\{i,t-1\}} + B_2 C_{\{j,t\}} + B_3 G_{\{j,t\}} + B_4 W_{\{j,t\}} + \\ + B_5 T_{\{j,t\}} + \lambda_t + \varepsilon_{\{i,t\}} \quad (1)$$

The dependent variable, $Q_{\{i,j,t\}}$ is a measure of water quality for lake *i* in HUC12 *j* at time *t*. For the measure of water quality, I use Carlson's Trophic State Index (CTSI). CTSI is

[3] http://limnology.eeob.iastate.edu/lakereport/default.aspx

[4] Although CDL data includes other crops such as wheat, they are more difficult to accurately identify. Data on CDL accuracy can be found at http://www.nass.usda.gov/research/Cropland/sarsfaqs2.htm

[5] Downloaded from http://prism.oregonstate.edu

an index of water quality from 0 to 100, where an increase indicates a degradation in water quality (Table 1).

Table 1: Summary statistics of the variables.

	mean	sd	10th percentile	90th percentile
water quality model:				
CTSI	60.81	9.89	47.00	73.00
Crops	5.17	2.89	2.05	8.98
Grass	2.26	1.51	0.63	4.26
Precipitation	39.26	16.60	21.21	58.96
Temperature	9.57	2.37	6.48	12.95
cropland model:				
Crops	5.41	2.99	1.74	9.61
Expected Price	1.70	0.55	1.00	2.51
Fertilizer	71.97	26.25	37.40	101.40
Fuel	68.73	23.45	33.40	99.30
Precipitation	38.77	16.33	22.20	55.75
Temperature	9.59	2.25	6.71	12.64
N	27,472			

Notes: The unit of observation for both models is a HUC12 watershed. The water quality model contains data from 2001-2016 for 123 Iowa lakes, with some gaps. CTSI stands for Carlson's Trophic State Index, a measure of water quality. The cropland model contains annual data from 2001-2016 for all 1717 HUC12 watersheds in Iowa. The sources for "Crops", "Precipitation", and "Temperature" are the same for both the water quality model and the cropland model.
Source: own estimations based on USDA and PRISM data

I include the lag in Q to account for the stock effects of lake water; the coefficient is expected to be positive as a certain amount of nutrients in a lake carry over across years. Including a lag of Q implies both short and long-term impacts from the other right-hand side variables on water quality. The short term, i.e. contemporaneous impacts are the estimated coefficients on the variables, while the long-term impacts are these coefficients multiplied by the dynamic multiplier (Greene, 2000):

$$\sigma = \frac{1}{(1 - B_1)} \tag{2}$$

The variables C and G represent year t, HUC12 j's hectares of cropland and grassland, respectively. The main coefficient of interest is , which measures the short-term marginal change in water quality due to an increase in cropland. Since an increase in CTSI represents lower water quality, is expected to be positive due to nutrient runoff. Of secondary interest is the marginal increase in water quality due to an increase in grassland, , which could be considered a proxy to the effect of CRP on water quality. A negative sign on would indicate beneficial qualities of increased grassland near lake water. Since the CRP program requires active cropland to be retired, the total short-term effect of CRP on lake water quality is $(-B_2 + B_3)$.

I control for the effect of weather on water quality by including annual measures of precipitation, W, and temperature T. Although the focus of the paper is on the effect of crops on water quality, the effect of weather on water quality is an important and complicated topic. For example, it is not clear a priori what the sign of these weather effects will be; increased rainfall, for example, can dilute existing nutrient levels, but can also increase nutrient runoff from nearby farms. Several papers have also highlighted the importance of studying the effects of weather on water quality, given the predicted increased variation in weather due to climate change (Delpla, 2009). The coefficient estimates on precipitation and temperature help shed light on these issues.

I control for time invariant, unobservable variables through lake level fixed effects, thus the coefficients are identified by the variation of the data within a lake. The unobservable variables could be, for example, geographic features that are fixed over time, such as soil quality, slope, or surface area, and it can also include permanent man-made structures that can alter the flow of water to lakes, such as tile drains. Year dummy variables control for unobserved trends over time.

Each HUC12 watershed is contained within larger watersheds which share drainage properties. To control for correlation between HUC12's within the same drainage area, I cluster standard errors at the HUC8 level.

Cropland Response Model

To estimate the response of cropland to crop prices I estimate the following Nerlovian partial adjustment model (Nerlove, 1956):

$$\begin{aligned} C_{\{j,t\}} &= \eta_j + B_1 C_{\{j,t-1\}} + B_2 E[P_{\{j,t\}}] + B_3 T_{\{j,t-1\}} + B_4 W_{\{j,t-1\}} + \\ &+ B_5 F_{\{t-1\}} + B_6 G_{\{t-1\}} + B_7 t + \varepsilon_{\{j,t\}} \end{aligned} \tag{3}$$

This model assumes that a representative farmer in watershed j make spring acreage decisions based on last year's acreage, climate and operating costs, as well as the expected crop prices during fall harvest. Operating costs consist of fertilizer, F, and fuel, G.

The variable of interest is which represents the marginal change in cropland due to an increase in expected prices. For the expected price I construct a Laspeyres price index using futures prices on corn and soybeans, along with observed soybean and corn acreages (Huang, 2010; Evans, 2015):

$$E[p_{\{j,t\}}] = \frac{\Sigma_c \left(E[P_{\{j,c,t\}} C_{\{j,c,t_0\}}]\right)}{\Sigma_c \left(E[P_{\{j,c,t_0\}} C_{\{j,c,t_0\}}]\right)} \tag{4}$$

where c is either corn or soybeans. Figure 2 shows the crop price index in Iowa from 2001-2016, averaged over HUC12 watersheds.

Estimation

Dynamic panel data models with fixed effects suffer from the well-known "Nickell bias", which results from the *within* transformation that subtracts the time mean from each group in order to remove the fixed-effects (Nickell, 1981). In a dynamic model, this will cause the lagged, transformed dependent variable to be correlated with the error term, violating the assumed orthogonality condition. One solution is to use the Arellano-Bond model (henceforth abbreviated as

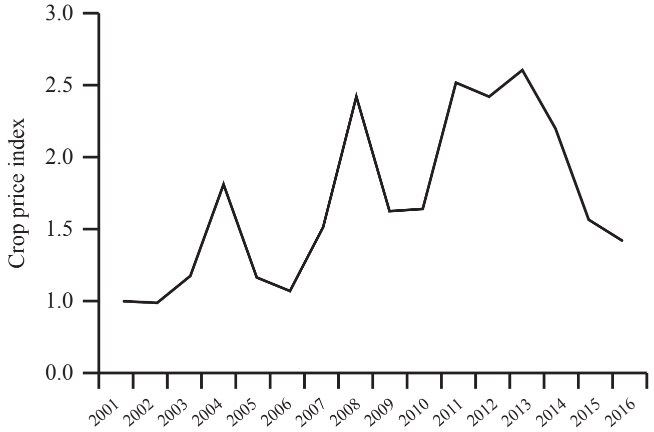

Figure 2: Crop Price Index for corn and soybeans in Iowa, averaged over HUC 12 watersheds from 2001-2016.

Note: Crop Price Index is a Laspeyres' Index using futures prices for corn and soybean along with corn and soybean acreage.
Source: Chicago Board of trade, retrieved from https://www.quandl.com/data/CHRIS-Wiki-Continuous-Futures

Table 2: Regression Results: Water Quality Model

	AB	OLS	FE
L.CTSI	0.2600***	0.7430***	0.3000*
	(0.0410)	(0.0261)	(0.0529)
Crops	0.0538***	0.0076*	0.0251
	(0.0179)	(0.0044)	(0.0252)
Grass	0.0032	-0.0085	-0.0044
	(0.0083)	(0.0053)	(0.0092)
Prec.	0.0460***	0.0312***	0.0406*
	(0.0139)	(0.0093)	(0.0201)
Temp.	0.0324	0.0307*	0.0185
	(0.0261)	(0.0157)	(0.0304)
N	1,484	1,607	1,607

Notes: Arellano-Bond (AB), OLS, and fixed effects (FE) coefficient estimates and standard errors. Each observation is a water quality measurement from a specific lake in Iowa. Data include the years 2001-2016, excluding 2008. All estimates include year fixed effects. CTSI stands for Carlson Trophic Secchi Index. Crops is equal to the sum of corn and soybean land. Standard errors are clustered by HUC8 watershed. *** stands for 1% of significance, ** for 5% and * for 10%
Source: own estimations based on USDA and PRISM data

AB), also known as the "difference GMM" (Arellano, 1991), which constructs instruments for the lagged dependent variable using transformations of the data.

The model's nickname comes from using first differences of the data to remove fixed effects. However, when there are gaps in the data, as is the case with the CTSI measurements, they can result in a significant loss of observations. For example, all lakes in the data set are missing the year 2008, so neither $\Delta Q_{\{i,2008\}}$, or $\Delta Q_{\{i,2009\}}$ can be included in the estimation. Instead of the first difference transformation, I employ the forward orthogonal deviations (FOD) transformation, where the mean of all future observations of a variable is subtracted from the current observation for each year. This purges the fixed effects and allows for more observations than the first differences in an unbalanced panel.

AB estimates are typically estimated in both one and two-step variants. The two-step model uses a weighting matrix that is the inverse of an estimate of Var(z'e), where *z* is the vector of instruments. This is the optimal weighting matrix in the sense that it is asymptotically efficient. However, in finite samples the two-step estimates have been shown to be biased downward. To fix this, I employ the finite sample bias correction described in (Windmeijer, 2005).

Results

Table 2 displays the results for the water quality model. All variables are estimated in log form, so the coefficients are interpreted as the elasticity of water quality with respect to each variable. Column (1) shows the estimated coefficients using the Arellano-Bond model, which instruments for the endogenous lagged dependent variable.

The coefficient on the lag of CTSI is statistically significant and positive, which provides evidence that water quality conditions persist over time. The coefficient on *Crops* is positive and statistically significant, indicating that an increase in cropland increases CTSI and therefore lowers water quality. This result is intuitive since an increase in cropland implies increased nutrients on the land, which can drain into neighboring water bodies. The elasticity for the *Crops* coefficient is 0.0538 in the short run and, using equation (1), 0.0727 in the long run. I do not find a statistically significant effect of grassland on lake water quality.

The coefficient on precipitation is positive and statistically significant. This indicates that the overall effect of precipitation on water quality is detrimental. In other words, the effect of runoff due to precipitation dominates the dilution effect of increased precipitation in lakes. I do not find evidence of an effect of temperature on lake water quality.

Columns (2) and (3) of table 2 display the results of OLS and fixed effects (FE) estimation for comparison. The OLS estimates may suffer from omitted variable bias since it does not control for time invariant fixed effects. The coefficient on the lag of *CTSI* is positive and statistically significant, and over twice as large in magnitude as the Arellano-Bond estimate. This positive bias is a direct result of the omitted variable bias, as shown in (Roodman, 2009). The coefficient on *Crops* is positive and statistically significant. The magnitude of the coefficient indicates that a 1% increase in cropland is associated with a 0.08% increase in CTSI (and therefore water quality is worse). This estimate is slightly higher than the

Arellano-Bond model. As with the Arellano-Bond model, the effect of precipitation is positive and significant, indicating that an increase in precipitation increases CTSI.

The fixed effects model, which does not instrument for the endogenous lagged variable, finds a positive and statistically significant coefficient on the lag of water quality. The magnitude of the coefficient, at 0.300, is like the Arellano-Bond model. The coefficients on *Crops* and *Grass* are not statistically significant. The coefficient on precipitation is positive and marginally significant, although the magnitude is like the other two models.

Table 3 displays the results for the cropland response model. Again, the variables are logged so that the estimates can be interpreted as the elasticity of cropland to a specific variable. All variables show a statistically significant effect on cropland. The variable of interest, price, shows the expected positive relationship with cropland. The magnitude of the elasticity of cropland to prices, 0.066, is small but comparable to other estimates from the literature (Barr, 2011; Evans, 2015). Using the dynamic multiplier, the long run elasticity is 0.104.

Table 3: Regression Results: Cropland Response Model.

	AB	OLS	FE
L.Crops	0.3040**	0.9510**	0.6700**
	(0.0798)	(0.0070)	(0.0278)
Price	0.0525**	0.0466*	0.0479*
	(0.0158)	(0.0222)	(0.0200)
L.lnPrec	0.0017	0.0235**	0.0160
	(0.0139)	(0.0084)	(0.0127)
L.lnTemp	-0.1670**	-0.0306**	-0.0798**
	(0.0263)	(0.0089)	(0.0151)
L.Fuel	-0.1060**	-0.1060**	-0.1010**
	(0.0224)	(0.0290)	(0.0273)
L.Fert	0.0446**	0.0828**	0.0684**
	(0.0112)	(0.0186)	(0.0147)
N	24,033	25,750	25,750

Notes: Arellano-Bond (AB), OLS, and fixed effects (FE) coefficient estimates and standard errors. Each observation is a HUC12 watershed in Iowa. Data include the years 2001-2016. Crops is equal to the sum of corn and soybean land. Standard errors are clustered by HUC8 watershed. ** stands for 1% of significance, * for 5%.
Source: own estimations based on USDA and PRISM data

The results for the control variables are mostly intuitive. The weather variable coefficients indicate that cropland decreases in response to increases in the previous year's temperature, while it increases in response to the previous year's precipitation. The magnitude of these responses is small and roughly equal. Increases in last year's fuel costs have a negative effect on this year's cropland. On the other hand, increases in the cost of last year's fertilizer have a positive effect on this year's cropland. Although this result is not intuitive, it has been found in other research (Evans, 2015; Huang, 2010).

Columns (2) and (3) show the OLS and fixed effects results of the cropland response model. The signs and magnitudes of the coefficients are similar across all three models, except for the coefficient on the lag of crops, where the FE and OLS estimates are larger than the Arellano-Bond estimate.

Application to the Renewable Fuel Standard

This section uses the previous elasticity estimates to measure the impact of the Renewable Fuel Standard (RFS) on lake water quality in Iowa. Since the RFS was enacted in 2005, I focus on the effects over 2006 to 2016, which is the most recent year with available data. The RFS mandated a large increase in ethanol, which is equivalent to a large increase in demand for corn since it is the primary feedstock. I follow the approach of Evans (2015) and use the price effects of this shock in demand to corn to connect the RFS to lake water quality. Specifically, I calculate the percent change in water quality using the following formula:

$$\%\Delta Q = \%\Delta Price\,\varepsilon_{\{wq,c\}}\varepsilon_{\{c,p\}} \quad (5)$$

where represents the elasticity of water quality with respect to cropland, and represents the elasticity of cropland to prices. I calculate (2) using both the short and long run elasticities, which can be considered lower and upper bounds.

For the change in price, I use estimates from (Hausman, 2012). Using a structural vector auto-regression (SVAR) model, the authors estimate that an increase in the demand of corn acreage for ethanol of .40 million hectares increases the price of corn and soybeans by 0.06 and 0.03 cents per cubic meter, respectively. According to data from USDA ERS, 1.29 billion cubic meters of corn were used to produce ethanol in 2005/2006, compared to 4.18 billion in 2015/2016, an increase of 2.89 billion cubic meters. Over the same period, the national average corn yield was 79 cubic meters of corn per hectare. Thus, the shock in demand is equivalent to an approximate 9.31 million hectare increase in demand for corn acreage. The above estimates from (Hausman, 2012) imply this increase in demand would increase the price of corn by $1.84 per bushel and the price of soybeans by $0.92 per bushel. I use these changes to calculate the counterfactual price index for each HUC12 in 2016. The average HUC12 experienced an approximate 58% increase in the price index because of the RFS.

Finally, I calculate the percentage change in lake water quality using both the short and long-term elasticities estimated from the previous analysis. The average lake experienced a 0.13% increase in CTSI in the short run, and a 0.27% increase in the long run. Figure 3 shows the distribution of these increases across the 123 lakes in the analysis. Thus, due to the very inelastic responses of lake water quality to cropland, and from cropland to crop prices, I find a negligible effect of the RFS on lake water quality.

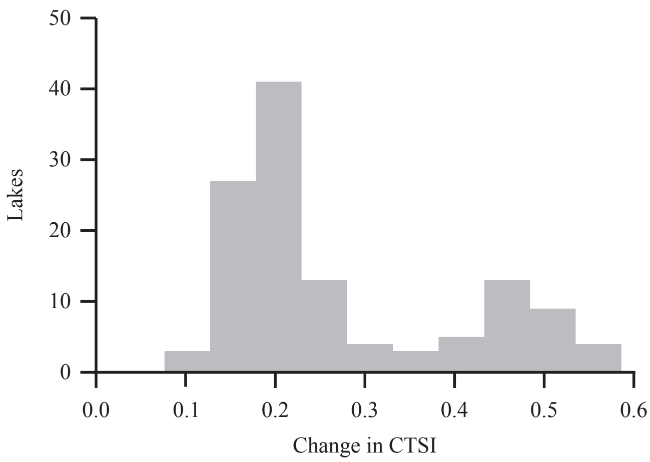

Figure 3: Estimated percentage change in CTSI due to the RFS.
Source: own estimates

Summary and Conclusions

U.S. agricultural and energy policies often have direct and indirect effects on the environment. Policies which affect agricultural land use can alter lake water quality through increased nutrient runoff. It is important to estimate these impacts to undertake thorough cost-benefit analyses of these policies.

This study focuses on estimating the effect of land use change on lake water quality in Iowa. High quality lake water measurements over 15 years are combined with satellite imagery and PRISM weather data to create a unique panel data set. Using a dynamic panel data model, I estimate the elasticity of CTSI to cropland to be 0.05% in the short run, and 0.07% in the long run, indicating that increases in cropland decrease lake water quality by a small amount. I also find a positive and significant coefficient on the lag of the dependent variable, which is evidence of a stock effect of lake water quality over time.

A second model estimated the elasticity of cropland to crop prices to be 0.066. Using these two elasticities, I estimate that the Renewable Fuel Standard decreased water quality by between 0.13 and 0.27%. The estimates may represent a lower bound since the paper only studies land use change along the extensive margin. Rather than expand cropland, farmers may alter crop rotations in favor of corn because of the RFS. Since corn requires a relatively high amount of fertilizer, the actual impact on water quality may be higher.

References

Arellano, M. and Bond, S. (1991): Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. The Review of Economic Studies 277-297. https://doi.org/10.2307/2297968

Atasoy, M., Palmquist, R.B. and Phaneuf, D.J. (2006). Estimating the effects of urban residential development on water quality using microdata. Journal of Environmental Quality 399-408. https://doi.org/10.1016/j.jenvman.2005.07.012

Balkcom, K.S. (2003): Testing soils and cornstalks to evaluate nitrogen management on the watershed scale. Journal of Environmental Quality, 1015-1024. https://doi.org/10.2134/jeq2003.1015

Barr, K.J. (2011): Agricultural land elasticities in the United States and Brazil. Applied Economic Perspectives and Policy, 449-462. https://doi.org/10.1093/aepp/ppr011

Dauer, D.M. (2000): Relationships between benthic community condition, water quality, sediment quality, nutrient loads, and land use patterns in Chesapeake Bay. Estuaries 80-96. https://doi.org/10.2307/1353227

Delpla, I. a.-V. (2009): Impacts of climate change on surface water quality in relation to drinking water production. Environment International, 1225-1233, https://doi.org/10.1016/j.envint.2009.07.001

Downing, J.A., Antoniou, G., Boatwright, L., Connrad, S., Kerdall, D., Li, J. and Gemesi, Z. (2006): Iowa Lakes Survey: Summer 2005 Scientific Initiative Reporting. Iowa Department of Natural Resources, Des Moines, Iowa. http://limnology.eeob.iastate.edu/lakereport/default.aspx

EPA, U. (2002): National Water Quality Inventory 2000 Report. United States Environmental Protection Agency Washington, D.C. https://www.epa.gov/waterdata/national-water-quality-inventory-report-congress

Evans, S.G. (2015): Effect of Agricultural Commodity Prices on Species Abundance of US Grassland Birds. Environmental and Resource Economics 549-565. https://doi.org/10.1007/s10640-014-9829-1

Greene, W.H. (2000): Econometric Analysis (5th edition). Prentice Hall. https://doi.org/10.1016/S0191-2615(02)00046-2

Hausman, C., Auffhammer, M. and Berck, P. (2012): Farm acreage shocks and crop prices: an SVAR approach to understanding the impacts of biofuels. Environmental and Resource Economics 117-136. https://doi.org/10.1007/s10640-012-9550-x

Huang, H. and Khanna, M. (2010): An econometric analysis of US crop yield and cropland acreage: implications for the impact of climate change. Working paper. http://ageconsearch.umn.edu/record/61527

Jeppesen, E. (2005): Lake responses to reduced nutrient loading--an analysis of contemporary long-term data from 35 case studies. Freshwater Biology 1747-1771. https://doi.org/10.1111/j.1365-2427.2005.01415.x

Lougheed, V.L.-F. (2001): Primary determinants of macrophyte community structure in 62 marshes across the Great Lakes basin: latitude, land use, and water quality effects. Canadian Journal of Fisheries and Aquatic Sciences 1603-1612. https://doi.org/10.1139/f01-102

Mankin, K.R.-H. (2003): Watershed-lake water quality modeling: Verification and application. Journal of Soil and Water Conservation 188-197. https://search.proquest.com/docview/220952274?accountid=960

Meador, M.R. and Goldstein, R.M. (2003): Assessing water quality at large geographic scales: relations among land use, water physicochemistry, riparian condition, and fish community structure. Environmental Management 0504-0517. https://doi.org/10.1007/s00267-002-2805-5

Nerlove, M. (1956): Estimates of the elasticities of supply of selected agricultural commodities. Journal of Farm Economics 496-509. http://www.jstor.org/stable/1234389

Nickell, S. (1981): Biases in dynamic models with fixed effects. Econometrica: Journal of the Econometric Society 1417-1426. https://doi.org/10.1016/0165-1765(88)90046-8

Rabotyagov, S.S., Kling, C.L., Gassman, P.W., Rabalais, N.N. and Turner, R.E. (2014): The Economics of Dead Zones: Causes, Impacts, Policy Challenges, and a Model of the Gulf of Mexico Hypoxic. Review of Environmental Economics and Policy 58-79. https://doi.org/10.1093/reep/ret024

Reckhow, R.H. (1984): Techniques for Detecting Trends in Lake Water Quality. Wiley Online Library. https://doi.org/10.1111/j.1752-1688.1984.tb04640.x

Roodman, D. (2009): How to do xtabond2: An introduction to difference and system GMM in Stata. Stata Journal 86. https://doi.org//10.2139/ssrn.982943

Schilling, K.E. (2014): The potential for agricultural land use change to reduce flood risk in a large watershed. Hydrological Research 3314-3325. https://doi.org/10.1002/hyp.9865

Secchi, S., Gassman, P.W., Williams, J.R. and Babcock, B.A. (2009): Corn-based ethanol production and environmental quality: A case of Iowa and the conservation reserve program. Environmental Management 732-744. https://doi.org/10.1007/s00267-009-9365-x

Simpson, T.W. (2008): The new gold rush: fueling ethanol production while protecting water quality. Journal of Environmental Quality 318-324. https://10.2134/jeq2007.0599

Sprague, L.A. (2012). Relating management practices and nutrient export in agricultural watersheds of the United States. Journal of Environmental Quality 1939-1950. https://10.2134/jeq2012.0073

Tong, S.T.Y. and Chen, W. (2002): Modeling the relationship between land use and surface water quality. Journal of Environmental Management 377-393. https://doi.org/10.1006/jema.2002.0593

Tu, J. (2011): Spatially varying relationships between land use and water quality across an urbanization gradient explored by geographically weighted regression. Applied Geography 376-392. https://doi.org/10.1016/j.apgeog.2010.08.001

Windmeijer, F. (2005): A finite sample correction for the variance of linear efficient two-step GMM estimators. Journal of Econometrics 25-51. https://doi.org/10.1016/j.jeconom.2004.02.005