# The Impact of Natural Amenity on Farmland Values: A Quantile Regression Approach.

**Hiroki Uematsu**
Graduate Research Assistant
Department of Agricultural Economics and Agribusiness
Louisiana State University
Room 28 Martin D. Woodin Hall
Baton Rouge, LA 70803
E-mail: huemat1@lsu.edu

**Ashok K. Mishra**
Professor
Department of Agricultural Economics and Agribusiness
Louisiana State University AgCenter
211 Martin D. Woodin Hall
Baton Rouge, LA 70803
Tel: 225-578-0262
Fax: 225-578-2716
E-mail: amishra@agcenter.lsu.edu

***Selected Paper prepared for presentation at the 2012 Southern Agricultural Economics Association Annual Meeting, Birmingham, AL, February 4-7, 2010***

# The Impact of Natural Amenity on Farmland Values: A Quantile Regression Approach.

Abstract

The objective of this study is to estimate the impact of natural amenity on farmland values in the contiguous United States using a quantile regression approach and data from the 2006, 2007, and 2008 Agricultural Resource Management Surveys. The contribution of this study is three-fold. First, we explicitly include variables representing natural amenity, and soil characteristics of farmland. Second, we employ a quantile regression approach to examine potentially heterogeneous impacts of natural amenity and soil characteristics at different quantiles of farmland values. Third, we utilized data from a nationwide survey of farm household to examine findings in studies using regional data are consistent at a national scale. Our quantile regression analysis offers some insightful results. Natural amenity is positively correlated with farmland values and its impact is often more pronounced at higher price range of farmland.

Keywords: Farmland Values, Quantile Regression, Natural Amenity

JEL Codes: C14, Q15, Q24

## I. Introduction

Land is a unique input in agriculture. Unlike other inputs, it is immovable, fixed in supply, and non-depreciable (Raup, 2003). Land is also one of the most important assets in agriculture, accounting for 75% to 85% of total assets in the U.S. farm sector (Gloy, et al., 2011, Sherrick and Barry, 2003). The well-being of farm households is critically dependent not only on earnings from farm operation, but also on farmland prices.

Identifying determinants of farmland values is, therefore, a classic topic in agricultural economics. A large number of empirical studies have been conducted to explore determinants of farmland value from a wide range of perspectives (Raup, 2003). The basic premise of the empirical studies is that current land price captures the net present value of future returns from the land (Guiling, et al., 2009). More recent empirical estimation of farmland values centers around the urban growth model developed by Cappoza and Helsley (1989) to incorporate the impact of urban development pressure on farmland values (Hardie, et al., 2001, Plantinga and Miller, 2001, Yue Jin, et al., 1997).

Despite the growing concern for natural resource conditions in rural America, only a few studies have utilized variables representing natural amenity in the estimation of farmland values (e.g., Bastian, et al., 2002, Ready and Abdalla, 2005). Bastian et al., (2002) found that wildlife habitat, angling opporunities, and scenic vistas contributes to higher land prices in Wyoming. Ready and Abdalla (2005), using data from Pennsilvania, discovered that agricultural open space increases nearby property values whereas a large-scale livestock operation has an opposite impact. These studies rely on regional data and the relationship between natural amenity and farmland prices has yet to be examined on a national scale, to the best of our knowledge.

Moreover, a vast majority of empirical studies on this topic employs a standard parametric model, OLS and its variants. Surprisingly, few published studies, if any, have used a more flexible semi-parametric regression model to explore potentially complex and heterogeneous relationship between farmland attributes and prices.

This study attempts to address the three issues. The objective of this study is to estimate the impact of natural amenity on farmland values by a semi-parametric quantile regression approach using data from Agricultural Resource Management Survey (ARMS) a nationwide survey of farm households conducted by the Economic Research Service and the National Agricultural Statistics Service.

The rest of the paper is organized as follows. The next section briefly introduces the hedonic pricing method, a common approach in estimating farmland values. Section III extensively discusses quantile regression approach and its advantage over the conventional OLS approach. Section IV discusses data used in this study, followed by empirical results in Section V. The final section offers some concluding remarks.

## II. Conceptual Framework

In the hedonic pricing method (HPM), price of a good can be explained by a vector of objectively measured attributes (Rosen, 1974). The hedonic price function can be represented as

$$P = P(\boldsymbol{z}), \tag{1}$$

where $\boldsymbol{z} = (z_1, z_2, \cdots, z_n)$ and $z_i, i = 1, \cdots n$, is individual attribute of the product (Palmquist, 1991). Consumers of farmland, in an attempt to maximize their utility, would choose a farmland parcel so that the marginal implicit price of the parcel with respect to $z_i$ is equal to the marginal rate of substitution between $z_i$ and wealth (Ready and Abdalla, 2005).
In the context of this study, $P$ is self-reported per acre price of farmland, obtained from the ARMS. $\boldsymbol{z}$ consists of two groups of land attributes: amenity attributes and other attributes. Following notation used in Bastian, et al. (2002), we denote amenity attributes as $\boldsymbol{z_A}$ and other attributes as $\boldsymbol{z_O}$.

## III. Empirical Framework

A common approach to estimating farmland values is to apply a standard parametric model, OLS and its variants, to equation (1). In such a model, conditional means of the dependent variable, i.e., farmland values, are estimated under a specific distributional assumption. One major disadvantage of this approach is that some attributes of farmland can have heterogeneous impacts on farmland values. Some attributes may be "luxury" and only impact farmland values that have a relatively higher price range while other attributes may be "necessity," affecting property values only at a lower price range.

This study attempts to overcome the disadvantages of OLS models by using a quantile regression approach originally developed by Koenker and Bassett (1978). In the rest of this section, we provide an overview of the theoretical background of quantile regression with an emphasis on its relative advantages over the conventional least square method.

### a. Quantile Regression

The fundamental objective of an econometric analysis is to delineate the true relationship between variables that are of interest to the researcher by making prediction about the population based on sample data. Econometricians strive to find a way to make their prediction as accurate as possible, or make their prediction errors as small as possible so that she can get as close as possible to the true relationship between the variables in the population (without rarely knowing what the true relationship really is).

One of the reasons for which there exists a wide variety of econometric tools to achieve the seemingly straightforward objective is due to various assumptions econometricians make about predictions and prediction errors. To make it clear the difference between the conventional least square method and quantile regression, we introduce the concept of loss function, which is a general theoretical framework from which one can derive both least square and quantile regression methods by applying different sets of assumptions.

Following Cameron and Trivedi (2005), loss function, $L$, is defined as

$$L(e) = L(y - \hat{y}), \tag{2}$$

where $e$ is the prediction error, $y$ is the dependent variable and $\hat{y}$ is the prediction of $y$. An important property of $L(e)$ is that it is increasing in $e$. The objective of an econometric analysis in this context comes down to minimization of the expected value of loss function, that is,

$$\min E[L(e)] = \min E[L(y - \hat{y})], \tag{3}$$

In choosing an appropriate estimator, a researcher has freedom in choosing the functional form of $L(e)$ and specification of $\hat{y}$. For every combination of the loss function and specification of the prediction, there exists a unique estimator. Least square estimator, for example, minimizes sum of squared errors in the sample and prediction is formed as a linear combination of a set of regressors, $x$, and estimated parameters, $\hat{\beta}$. In the context of the loss function, least square estimator can be expressed as

$$\min E(e^2) = \min E\left(y - x'\hat{\beta}\right)^2, \tag{4}$$

In fact, this particular type of loss function is known as the squared error loss function for which the optimal $\hat{y}$ is the conditional mean function, $E(y|x)$ (Cameron and Trivedi, 2005). Least square estimator is a special case of the squared error loss function when $\hat{y}$ is assumed to be a linear function in $x$ and $\hat{\beta}$.

Least square estimator is also called a pure location shift model (Heckman, 1979) in the sense that it only estimates conditional means of $y$ given $x$. A very restrictive assumption in the location shift model is that conditional distributions of $y$ are identical at any values of $x$, except for the means, which are to be estimated by the least square method[1]. For all conditional distributions of $y$ given $x$, variances, skewness and kurtosis are assumed to be identical. Therefore, only change in the conditional distribution of $y$ due to change in $x$ is its relative location, which is determined by the conditional mean. Hence a location shift model.

As well articulated by Mosteller and Tukey (2001), the regression curves is a grand summary of the means of the conditional distributions and it gives an incomplete picture for a set of distributions for the same reason that the mean gives an incomplete picture of a single distribution. In reality, it may well be the case that conditional distribution of $y$ can be skewed or fat-tailed; there is no guarantee that conditional distribution of $y$ will always be unimodal. Despite such restrictive and naïve assumptions, most of applied econometric analyses concern with the conditional means (Fitzenberger, et al., 2002). It is these limitations in the location shift model that quantile regression can overcome by taking different functional form of the loss function.

Quantile regression is originally proposed by Koenker and Bassett (1978). Instead of estimating conditional means, $E(y|x)$, quantile regression can estimate any points on the conditional distribution by estimating conditional quantiles, $Q(\beta_q)$[2]. That is, $q$th quantile regression estimator is the one that minimizes the following objective function

---

[1] Of course, least squares method can be extended to Generalized Least Squares method to handle heterogeneous variances in conditional distributions, i.e, heteroskedasticity.

[2] Quantiles are to percentiles what probabilities are to percentages. For example, the 0.50 quantile is the 50th percentile Cameron and Trivedi (2009).

$$Q(\beta_q) = \min_{\beta \in R^p} \left[ \sum_{i \in \{i: y_i \geq x_i' \beta\}}^{N} q\left|y_i - x_i'\beta_q\right| + \sum_{i \in \{i: y_i < x_i' \beta\}}^{N} (1-q)\left|y_i - x_i'\beta_q\right| \right], \qquad (5)$$

$$q \in (0,1)$$

where $q$ is an arbitrarily chosen quantile and $p$ is the number of parameters to be estimated (Gould, 1993, Koenker and Bassett, 1978). In the context of the loss function, quantile regression minimizes a weighted sum of absolute values of errors with different weights being placed on positive and negative errors (Koenker and Hallock, 2001). Another advantage of quantile regression evident in (5) is that it is more robust to outliers as quantile regression estimates conditional quantiles instead of conditional means. The objective function (5) is not differentiable and thus the usual Newton Raphson algorithm cannot be used. The minimization problem can be solved by linear programming using Simplex method(Cameron and Trivedi, 2005, Koenker and Bassett, 1978). The asymptotic distribution of quantile regression can be shown as

$$\widehat{\beta_q} \sim N\left(\beta_q, \quad \boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{A}^{-1}\right), \qquad (6)$$

where $\boldsymbol{A} = \sum_i q(1-q)x_i\, x_i{}'$, $\boldsymbol{B} = \sum_i f_{u_q}(0|x_i)\, x_i x_i{}'$ and $f_{u_q}(0|x_i)$ is the conditional density function of the error term for $q$th conditional quantile , $u_q = y - x'\widehat{\beta_q}$ evaluated at $u_q = 0$(Gould, 1993).

The analytic formula in (6) was originally developed by Koenker and Basset (1978) and it assumes that the error distribution is homoscedastic (Abdulai and Huffman, 2005). Bootstrap method originally proposed by Efron (2000) is more desirable as it allows us to obtain standard error estimates without any assumptions, considering the fact that the primary advantage of quantile regression is to avoid any distributional assumptions (Abdulai and Huffman, 2005, Cade and Noon, 2003). Simulation results by Rogers (2005) and Gould (2003) confirm this view; they demonstrated that, under heteroskedasticity, standard error estimates based on the analytic formula understated true standard errors and bootstrap method obtained more accurate estimates. Koenker and Hallock (1979) estimated a log wage model with quantile regression estimator and obtained bootstrap standard errors with 20, 200 and 600 replications. They found that estimates with 600 replication to be more stable, as suggested by Andrews and Buchinsky (2000).

To summarize the discussion so far, we list major advantages of quantile regression over the conventional least square method following Buchinsky (1998). First, quantile regression provides a more complete picture of the conditional distributions of the dependent variable given a set of regressors; researchers can estimate any point on the conditional distributions and as many points on the conditional distribution as they wish to estimate. Different coefficient estimates at different quantiles would be a manifestation that a pure location model is inadequate to explain the underlying relationship between the variables of interest. Second, linear programming makes estimation of quantile regression is relatively easy. Third, the estimated coefficient vector of quantile regression is more robust to outliers as the objective function minimizes the weighted sum of absolute deviations. Fourth, quantile regression can be more efficient than least square method when error term is non-normal. Although computational cost of calculating bootstrap standard errors had been a major drawback of quantile regression, it is no longer the case with the recent advancement in the computer processing speed.

b. Empirical Application of Quantile Regression

Despite its applications in a wide range of topics in economics, quantile regression is less frequently used in agricultural economics. Within the limited amount of empirical work, the rest of the section will review some recent empirical applications of quantile regression that are fairly closely related to the topics of farmland value estimation.

Zietz, et al. (2008) employed spatial quantile regression model to address the major disagreement in the real estate literature over the direction and magnitude of the effect of certain characteristics on housing prices. Some interesting results from the quantile regression hedonic pricing model are that (1) premium for newness is the highest for the lower price ranges, (2) the number of bedroom has a significant and positive impact on the price only in lower and middle priced houses, indicating higher marginal value of an extra bedroom in the lower priced houses, (3) the additional bathroom has a higher marginal value in the higher priced houses and (4) marginal value of square footage for the higher priced houses is almost two and a half times larger than that for lower priced houses.

Heintzelman (2010) also employed quantile regression to estimate the impact on home prices of the passage of a Massachusetts policy that raise funds for historic preservation and affordable housing through referenda. The study observed that, in one of the two areas studied, the passage of such a referendum has no impact on house price below the 0.65 quartile, but a positive and significant impact is found for house prices at the upper quartiles. The author elaborated that this may be because owners of higher price houses have stronger preferences for preservation despite an additional tax burden.

Kostov (2009) estimated a spatial lag quantile regression model to estimate agricultural land prices in Northern Ireland. An interesting finding is that coefficient estimates at the top quantiles ($q > 0.95$) are markedly different from those at the rest of the sample. The test for equal coefficient estimates across different quantiles developed by Koenker and Bassett (1982) was rejected and the author argued that this is an evidence of two separate segments in agricultural land market in Northern Ireland, where good quality agricultural land are scarce.

What is common among all of the studies reviewed above is that, by using quantile regression, they have yielded additional insight into the conditional distribution of the dependent variable that would otherwise have been left unnoticed had they employed the traditional least square approach. Although empirical application of quantile regression in agricultural economics have been limited so far, the findings from the studies we reviewed above assures potential of quantile regression model in estimating the hedonic pricing model of farmland values.

IV. Data

The study employs data obtained from the nationwide Agricultural Resource Management Survey (ARMS) from 2006 to 2008, developed by the Economic Research Service and the National Agricultural Statistical Service. The 2006-2008 ARMS surveys provide information about the relationships between farmland values, agricultural production, resources, and the environment as well as about the characteristics and financial conditions of farm households. Data are collected from one operator per farm, the senior farm operator, who makes most of the day-to-day management decisions. We also utilize county-level variables developed by the United States Department of Agriculture: the Amenity Index and its components to represent natural amenity of farmland, average monthly rainfall data between 1960 and 2003 to represent production uncertainty, soil productivity index to measure land quality, the 2000 Population Interaction Zones in Agriculture (PIZA) that captures intensity of population interaction. As in Guiling, et al. (2009), the dataset used in this study is unique in that it supplements parcel specific data with county-level variables to estimate a hedonic pricing model of farmland values.

The amenity index is developed by McGranahan (1999). It captures environmental attributes of farmland such as climate, topography, and water area that are highly correlated with farmland values (McGranahan, 1999). The amenity index is an ordinal variable ranging from 1 to 7 with 1 being the lowest amenity and 7 being the highest amenity. It consists of 1) water area as proportion of total county area, 2) average temperature in January, 3) average temperature in July, 4) mean hours of sunlight in January, 5) average humidity in July, 6) land surface form topography codes. For more details about the amenity index, see McGranahan (1999).

In our empirical analysis, we estimate two versions of quantile regression model, corresponding to equation (1). In the first model, we utilize the amenity index as an aggregate measure of natural amenity. Then we estimate the same model using each of the six individual components of the amenity index[3] to examine if any additional insights can be gleaned by decomposing the index.

Other regressors are selected following existing studies and theoretical expectations. To supplement the amenity index, we include county-level average monthly rainfall and standard deviation and soil productivity index also aggregated at the county level. To control for the extensively studied impact of government payments on land values, we include direct payments, indirect payments, disaster payments, land retirement payments, and working land conservation payments, all of which are calculated on per acre basis. We also include logarithm of value of production per acre and dummy variables for farms whose primary enterprise is livestock or high value crops. The effect of urban sprawl is controlled by Population Interaction Zone in Agriculture in 2000 (PIZA) and dummy variables for farms located in metro county and rural county. Since we pool three years of ARMS data, dummy variables for observation years 2006 and 2007 are also included. Finally, regional dummy variables are used to capture local land characteristics that cannot be captured by all other variables[4]. A total of 15,108 observations are used in this study. Table 1 provides variable definitions and summary statistics.

## V. Empirical Results

Table 2 presents results from first of the two quantile regression model in which the amenity index is used as an aggregate measure of land amenity. Table 3 provides results from the other model that uses the six individual components of the amenity index. In both tables, the second column shows OLS coefficient estimates, for the sake of comparison. The

---

[3] A total of 14 variables are used, including some squared terms and dummy variables for topological classifications.

[4] See Figure 1 for a map of USDA Farm Resource Regions.

third through seventh columns display quantile regression coefficient estimates. The last column shows F-statistics from the Wald test that examines if at least one coefficient is significantly different from other coefficients estimated at other quantiles. A significant F-score underlines the suitability of quantile regression approach over the conventional OLS approach.

The primary interest of this study is the relationship between natural amenity and farmland values. In Table 2, the amenity index has a positive and significant coefficient at all quantiles and the coefficient estimates become larger at higher quantiles. The Wald F-score of 11.59 indicates that the impact of the amenity index on farmland values is different at different conditional quantiles of farmland values. The finding here is consistent with our expectation that natural amenity is more "luxury" than "necessity" and that it has a larger impact on farmland values at a higher price range.

In the other model (Table 3), a total of 14 variables are included in place of the amenity index to further investigate the impact of natural amenity on farmland values. Compared to OLS coefficient estimates, quantile regression results offer a very different picture of natural amenity and farmland values. Percent of water area has a positive impact on farmland values, but only at higher quantiles (0.50, 0.75, 0.90). Coastal areas and areas with lakes can be more scenic and pleasing (McGranahan, 1999) and being in proximity of a body of water can be a luxury attribute of land and thus it is affecting farmland values only at higher quantiles. Although the OLS estimate is also positive and significant, the additional insights can only be obtained with quantile regression. Mean temperature in July has a positive impact at 0.25 quantile while, at the highest quantile, it has a negative impact. It is conceivable that a high average temperature in the summer can be detrimental to farmland values at the very high price range. Average relative humidity has no impact on farmland values except at the lowest quantile (0.10) where it has a positive and significant impact on farmland values. This may indicate the importance of humidity for agricultural production as a basic necessity for land parcels at the very low price range. Topological classification yielded somewhat unexpected results. Relative to the base group of "open hill or mountains," land parcels classified as "Plains" are negatively correlated with farmland values at all quantiles and "Tablelands" have negative coefficients only at the lower quantiles (0.10 and 0.25). While "Plains with Hills or Mountains" have no significant impact, "Hills or Mountains" is positively correlated with farmland values at all quantiles. Using individual components of the amenity index (Table 3) allows us to observe a more detailed picture of how farmland values are influenced through different modules of natural amenity.

In both models, the rest of the variables have obtained very similar results with one exception. That is, in Table 2, the mean soil productivity index is unexpectedly negatively correlated with farmland values at lower quantiles (0.10, 0.25, and 0.50) whereas, in Table 3, it is positively correlated with farmland values at all but the lowest quantile and the impact becomes larger at higher quantiles. The later result is consistent with the intuitive expectation that more productive land parcels are more expensive. Note that the only difference between the two models is the way in which natural amenity is represented. In the first model (Table 2), the amenity index is used as an aggregate measure of natural amenity whereas in the other model (Table 3), a total of 14 variables are included to capture various aspects of natural amenity. Thus, a plausible explanation to the unexpected result found in Table 2 is due to the omitted variable problem; since the second model

(Table 3) captures a wide range of natural amenity factors that are aggregated into one index in the first model (Table 2), it can more accurately represent the association between soil productivity and farmland values.

To our surprises, a majority of government payments have no impact on farmland values in both models. This may because we have broken down government payments into five different categories instead of having one aggregate variable representing total government payment received. Alternatively, this could be due to the fact that we have introduced variables representing natural amenity that are absent in many existing studies of income capitalization models. The unexpected result here calls for further research.

Farm Variables obtained expected results in both models. Value of production per acre is positively correlated with farmland values at all quantiles in both models. Land parcels owned by a farm whose primary enterprise is livestock negatively affects farmland values and the negative impact is stronger at lower price ranges (Tables 2 and 3). Ready and Abdalla (2005) confirmed the negative impact of a large livestock operation on nearby property values in Pennsylvania, and the results here confirms the same effect on a national scale. Land parcels owned by high-value crop farms, on the other hand, is positively correlated with farmland values except at the lowest quantiles. In both models, the effect is higher at higher quantiles, indicating that production of high value crops such as fruits and vegetables is a luxury attribute to farmland values.

Population interaction in county in which farm is located positively influences farmland values, due to urban development pressure on farmland values (Hardie, et al., 2001, Plantinga and Miller, 2001, Yue Jin, et al., 1997). Likewise, land parcels located in metro county is more expensive and those located in rural counties are less expensive. Dummy variables for observation years shows expected results. Farmland values are evaluated higher in 2006 than in 2007 or 2008, presumably due to the economic recession that have plagued the economy since 2007.

Finally, regional dummy variables have obtained significant results in both models. Compared to the base group of the Heartland region comprising of major agricultural states in the Midwest, all the regions are negatively correlated with farmland values, with a few exceptions. The first exception is the Fruitful Rim region, mainly comprising of coastal areas including Florida and California (Figure 1). The dummy variable for the Fruitful Rim region is positive and significant at quantiles higher than 0.50 in both models. Another exception is the Basin and Range region that has a negative and significant impact at 0.50 quantile or lower but a positive and significant impact at 0.75 and 0.90 quantiles in both models. The fact that a vast majority of regional dummy variables have significant coefficient indicates that there are some important underlying variables that have significant explanatory power but are unavailable to the researcher. Exploring new variables that have yet to be used in existing studies is warranted to refine our empirical findings.

## VI. Conclusion

The objective of this study is to estimate the impact of natural amenity on farmland values in the contiguous United States using a quantile regression approach and data from the 2006, 2007, and 2008 Agricultural Resource Management Surveys. The contribution of this study is three-fold. First, we explicitly include variables representing natural amenity, and soil characteristics of farmland. Second, we employ a quantile regression approach to examine potentially heterogeneous impacts of natural amenity and soil characteristics at different quantiles of farmland values. Third, we utilized data from a nationwide survey of farm household to examine findings in studies using regional data are consistent at a national scale.

Our quantile regression analysis offers some insightful results. Natural amenity is positively correlated with farmland values and its impact is often more pronounced at higher price range of farmland. Some attributes such as water area as proportion of total county area and high value crop farms are "luxury" in that they increase farmland values only at higher quantiles. On the other hand, average humidity has a positive and significant impact on farmland values at the lower quantiles (0.10 and 0.25), indicating that humidity is a "necessity."

Evaluation of farmland values have been a major policy issues since the 1980s due to the growing urban development pressure (Livanis, et al., 2006). This paper contributes to the existing literature of farmland valuation by explicitly incorporating natural amenity and soil characteristics of farmland and employing a semi-parametric quantile regression model.

References:


Abdulai, A., and W.E. Huffman. 2005. "The Diffusion of New Agricultural Technologies: The Case of Crossbred-Cow Technology in Tanzania." *American Journal of Agricultural Economics* 87:645-659.

Bastian, C.T., D.M. McLeod, M.J. Germino, W.A. Reiners, and B.J. Blasko. 2002. "Environmental amenities and agricultural land values: a hedonic model using geographic information systems data." *Ecological Economics* 40(3):337-349.

Buchinsky, M. 1998. "Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research." *The Journal of Human Resources* 33(1):88-126.

Cade, B.S., and B.R. Noon. 2003. "A gentle introduction to quantile regression for ecologists." *Frontiers in Ecology and the Environment* 1(8):412-420.

Cameron, A.C., and P.K. Trivedi. 2005. *Microeconometrics Methods and Applications*: Cambridge University Press.

Capozza, D.R., and R.W. Helsley. 1989. "The fundamentals of land prices and urban growth." *Journal of Urban Economics* 26(3):295-306.

Fitzenberger, B., R. Koenker, and J.A.F. Machado, eds. 2002. *Economic Applications of Quantile Regression*: Physica-Verlag HD.

Gloy, B.A., M.D. Boehlje, C.L. Dobbins, C. Hurt, and T.G. Baker. 2011. "Are Economic Fundamentals Driving Farmland Values?" *Choices* 26(2).

Gould, W. 1993. "Quantile regression with bootstrapped standard errors." *Stata Technical Bulletin* 2(9).

Guiling, P., B.W. Brorsen, and D. Doye. 2009. "Effect of Urban Proximity on Agricultural Land Values." *Land Economics* 85(2):252-264.

Hardie, I.W., T.A. Narayan, and B.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(1):120-132.

Heckman, J.J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47(1):153-161.

Heintzelman, M.D. 2010. "Measuring the Property-Value Effects of Local Land Use and Preservation Referenda." *Land Economics* 86:22-47.

Huffman, W.E. (2001) Human capital: Education and agriculture, ed. B.L. Gardner, and G.C. Rausser, vol. 1A, Elsevier, pp. 333-381.

Koenker, R., and G. Bassett. 1982. "Tests of Linear Hypotheses and l"1 Estimation." *Econometrica* 50(6):1577-1583.

Koenker, R., and G. Bassett, Jr. 1978. "Regression Quantiles." *Econometrica* 46(1):33-50.

Koenker, R., and K.F. Hallock. 2001. "Quantile Regression." *Journal of Economic Perspectives* 15(4):143-156.

Koenker, R., and K.F. Hallock (2000) Quantile Regression An Introduction.

Kostov, P. (2009) A Spatial Quantile Regression Hedonic Model of Agricultural Land Prices, vol. 4, pp. 53-72.

Livanis, G., C.B. Moss, V.E. Breneman, and R.F. Nehring. 2006. "Urban Sprawl and Farmland Prices." *American Journal of Agricultural Economics* 88(4):915-929.

McGranahan, D.A. 1999. "Natural Amenities Drive Rural Populaiton Change." *Agricultural Economic Report No. 781. Economic Research Service, U.S. Department of Agriculture.*

Palmquist, R.B. (1991) Hedonic Methods, ed. J.B. Braden, and C.D. Kolstad, North-Holland, pp. 77-120.
Plantinga, A.J., and D.J. Miller. 2001. "Agricultural Land Values and the Value of Rights to Future Land Development." *Land Economics* 77(1):56-67.
Raup, P.M. (2003) Disaggregating Farmland Markets, ed. C.B. Moss, and A. Schmitz, Iowa State Press, pp. 15-25.
Ready, R.C., and C.W. Abdalla. 2005. "The Amenity and Disamenity Impacts of Agriculture: Estimates from a Hedonic Pricing Model." *American Journal of Agricultural Economics* 87(2):314-326.
Rosen, S. 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy* 82(1):34-55.
Sherrick, B.J., and P.J. Barry (2003) Farmland Markets: Historical Perspectives and Contemporary Issues, ed. C.B. Moss, and A. Schmitz, Iowa State Press, pp. 27-49.
Yue Jin, S., T.T. Phipps, and D. Colyer. 1997. "Agricultural land values under urbanizing influences." *Land Economics* 73(1):90.
Zietz, J., E. Zietz, and G. Sirmans. 2008. "Determinants of House Prices: A Quantile Regression Approach." *The Journal of Real Estate Finance and Economics* 37(4):317-333.

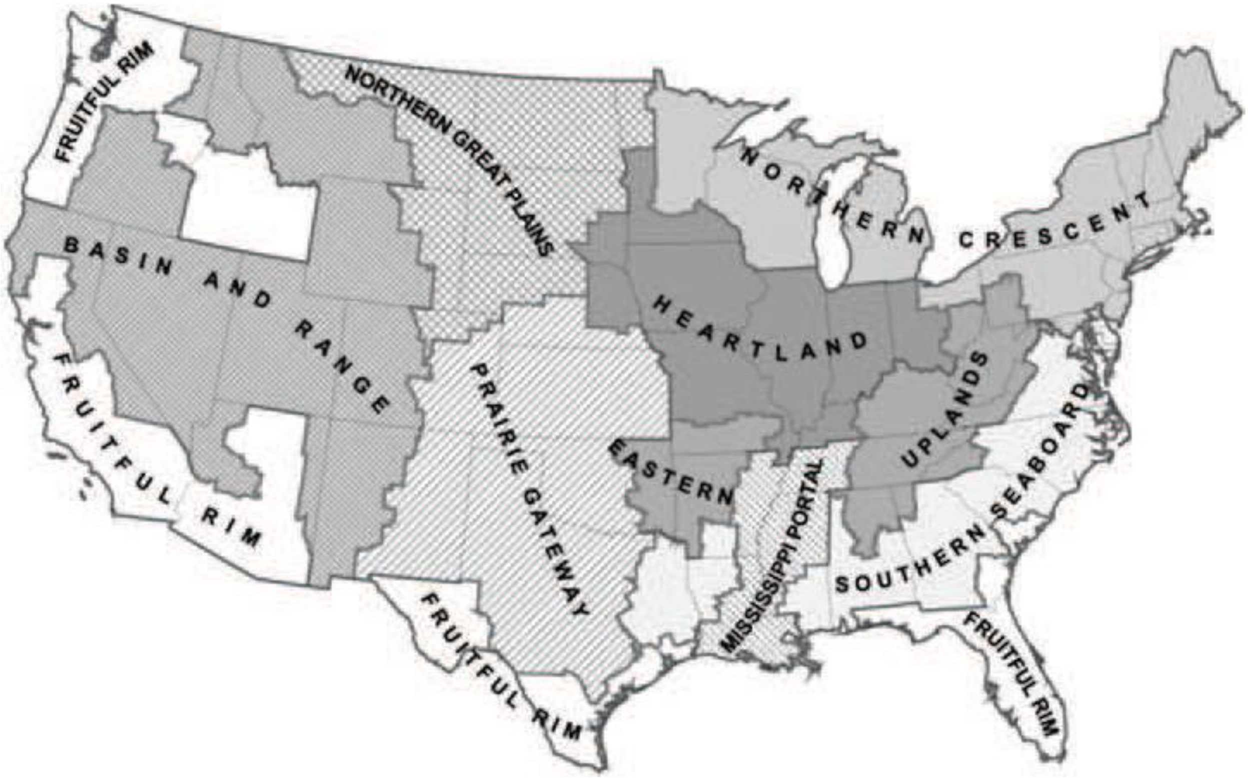


Figure 1: USDA Farm Resource Regions
Source: USDA ERS

Table 1: Variable Definitions and Summary Statistics

| Variables | Mean | Std. Dev | Min | Max |
|---|---|---|---|---|
| Amenity Index | 3.54 | 1.13 | 1 | 7 |
| Components of Amenity Index | | | | |
| Percent Water Area in County | 4.25 | 10.25 | 0.00 | 75.00 |
| Mean Temperature for January 1941-1970 | 33.39 | 12.44 | 1.10 | 66.80 |
| Mean Temperature for July 1941-1970 | 75.88 | 5.29 | 55.60 | 93.70 |
| Mean Hours of Sunlight 1941-1970 | 149.71 | 35.57 | 48.00 | 266.00 |
| Mean Relative Humidity 1941-1970 | 54.84 | 15.72 | 18.00 | 80.00 |
| Plains (=1 if yes) | 0.18 | 0.39 | 0.00 | 1.00 |
| Tablelands (=1 if yes) | 0.06 | 0.24 | 0.00 | 1.00 |
| Plains with Hills or Mountains(=1 if yes) | 0.08 | 0.27 | 0.00 | 1.00 |
| Hills or Mountains (=1 if yes) | 0.14 | 0.34 | 0.00 | 1.00 |
| County Level Variables | | | | |
| Mean of County-level Monthly Rainfall | 953.54 | 372.60 | 86.74 | 2,895.28 |
| Standard Deviation of County-level Monthly Rainfall | 57.42 | 111.57 | 1.58 | 916.29 |
| Mean Soil Productivity Index | 71.29 | 14.45 | 24.13 | 96.16 |
| Government Payments | | | | |
| Direct Payments per Owned Acre | 214.74 | 3,483.91 | 0.00 | 247,500.00 |
| Indirect Payments per Owned Acre | 25.42 | 697.01 | 0.00 | 44,691.00 |
| Disaster Payments per Owned Acre | 26.95 | 1,172.83 | 0.00 | 99,721.00 |
| Land Retirement Payments per Owned Acre | 13.80 | 618.68 | 0.00 | 62,993.00 |
| Working Land Conservation Payments per Owned Acre | 37.01 | 1,681.50 | 0.00 | 150,000.00 |
| Farm Variables | | | | |
| Value of Production per Acre (Log) | 6.19 | 2.21 | -4.05 | 17.00 |
| Livestock Farms (=1 if yes) | 0.53 | 0.50 | 0.00 | 1.00 |
| High Value Crop Farms (=1 if yes) | 0.14 | 0.35 | 0.00 | 1.00 |
| Other Variables | | | | |
| Population Interaction Zone in Agriculture | 1.60 | 0.91 | 1.00 | 4.00 |
| Farms located in Metro County(=1 if yes) | 0.40 | 0.49 | 0.00 | 1.00 |
| Farms located in Rural County(=1 if yes) | 0.08 | 0.27 | 0.00 | 1.00 |
| Year 2006 (=1 if observation in 2006) | 0.33 | 0.47 | 0.00 | 1.00 |
| Year 2007 (=1 if observation in 2007) | 0.32 | 0.47 | 0.00 | 1.00 |
| Regional Dummy Variables (Heartland Region is excluded) | | | | |
| Northern Crescent Region | 0.15 | 0.36 | 0.00 | 1.00 |
| Northern Great Plains Region | 0.05 | 0.22 | 0.00 | 1.00 |
| Prairie Gateway Region | 0.11 | 0.31 | 0.00 | 1.00 |
| Eastern Uplands Region | 0.11 | 0.31 | 0.00 | 1.00 |
| Southern Seaboard Region | 0.15 | 0.35 | 0.00 | 1.00 |
| Fruitful Rim Region | 0.17 | 0.37 | 0.00 | 1.00 |
| Basin and Range Region | 0.05 | 0.21 | 0.00 | 1.00 |
| Mississippi Portal Region | 0.06 | 0.23 | 0.00 | 1.00 |
| Number of Observations | | 15,108 | | |

Table 2: Quantile Regression with Amenity Index

| Variables | OLS | Estimated Quantile | | | | | Wald |
|---|---|---|---|---|---|---|---|
| | | 0.10 | 0.25 | 0.50 | 0.75 | 0.90 | F-score |
| Amenity Index | 0.152*** | 0.047** | 0.055*** | 0.091*** | 0.123*** | 0.200*** | 11.59*** |
| County Level Variables | | | | | | | |
| Mean of County-level Monthly Rainfall | 0.001*** | 0.001*** | 0.001*** | 0.001*** | 0.001*** | 0.000 | 3.06** |
| Mean of County-level Monthly Rainfall Squared | 0.000*** | 0.000*** | 0.000*** | 0.000*** | 0.000*** | 0.000* | 1.56 |
| Standard Deviation of County-level Monthly Rainfall | 0.000 | 0.000 | 0.000 | 0.000 | 0.000** | 0.001*** | 7.48*** |
| Mean Soil Productivity Index | -0.002*** | -0.003** | -0.002* | -0.001* | 0.000 | -0.001 | 1.26 |
| Government Payments | | | | | | | |
| Direct Payments per Owned Acre | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.82 |
| Indirect Payments per Owned Acre | 0.000 | 0.000 | 0.000 | 0.000 | 0.000** | 0.000 | 0.89 |
| Disaster Payments per Owned Acre | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.1 |
| Land Retirement Payments per Owned Acre | 0.000*** | 0.000 | 0.000 | 0.000* | 0.000* | 0.000 | 0.09 |
| Working Land Conservation Payments per Owned Acre | 0.000*** | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.87 |
| Farm Variables | | | | | | | |
| Value of Production per Operated Acre (Log) | 0.103*** | 0.087*** | 0.093*** | 0.094*** | 0.098*** | 0.089*** | 1.08 |
| Livestock Farms (=1 if yes) | -0.101*** | -0.212*** | -0.166*** | -0.093*** | -0.053*** | -0.034*** | 9.02*** |
| High Value Crop Farms (=1 if yes) | 0.188*** | 0.025 | 0.100*** | 0.177*** | 0.322*** | 0.457*** | 12.06*** |
| Other Variables | | | | | | | |
| Population Interaction Zone in Agriculture | 0.292*** | 0.226*** | 0.225*** | 0.266*** | 0.339*** | 0.418*** | 22.75*** |
| Farms located in Metro County(=1 if yes) | 0.166*** | -0.031 | 0.094*** | 0.148*** | 0.165*** | 0.165*** | 12.11*** |
| Farms located in Rural County(=1 if yes) | -0.153*** | -0.080** | -0.093*** | -0.161*** | -0.214*** | -0.209*** | 3.5*** |
| Year 2006 (=1 if observation in 2006) | 0.051*** | 0.082*** | 0.064*** | 0.066*** | 0.059*** | 0.058* | 0.25 |
| Year 2007 (=1 if observation in 2007) | 0.015 | 0.026 | 0.020 | 0.031* | 0.042** | 0.054* | 0.25 |
| Intercept | 6.139*** | 5.623*** | 5.949*** | 6.324*** | 6.396*** | 6.802*** | |
| Regional Dummy Variables (Heartland Region is excluded) | | | | | | | |
| Northern Crescent Region | -0.291*** | -0.496*** | -0.344*** | -0.228*** | -0.208*** | -0.182*** | 11.19*** |
| Northern Great Plains Region | -1.047*** | -1.113*** | -1.148*** | -1.128*** | -0.866*** | -0.862*** | 6.72*** |
| Prairie Gateway Region | -0.839*** | -0.911*** | -0.900*** | -0.823*** | -0.727*** | -0.693*** | 3.52*** |
| Eastern Uplands Region | -0.377*** | -0.425*** | -0.480*** | -0.392*** | -0.310*** | -0.150*** | 8.5*** |
| Southern Seaboard Region | -0.254*** | -0.444*** | -0.346*** | -0.261*** | -0.134*** | 0.038 | 11.75*** |
| Fruitful Rim Region | 0.030 | -0.302*** | -0.029 | 0.202*** | 0.322*** | 0.348*** | 11.88*** |
| Basin and Range Region | -0.257*** | -0.703*** | -0.585*** | -0.265*** | 0.185* | 0.366** | 17.65*** |
| Mississippi Portal Region | -0.535*** | -0.561*** | -0.585*** | -0.551*** | -0.545*** | -0.453*** | 1.42 |
| Pseudo $R^2$ ($R^2$ for OLS) | 0.386 | 0.220 | 0.230 | 0.230 | 0.254 | 0.298 | |
| Number of Observations | 15,108 | | | | | | |

***, **, and * indicate statistical significance at 1%, 5%, and 10% levels, respectively.

Table 3: Quantile Regression with Components of Amenity Index

| Variables | OLS | Estimated Quantile | | | | | Wald F-score |
|---|---|---|---|---|---|---|---|
| | | 0.10 | 0.25 | 0.50 | 0.75 | 0.90 | |
| **Components of Amenity Index** | | | | | | | |
| Percent Water Area in County | 0.011*** | 0.002 | 0.003 | 0.008*** | 0.013*** | 0.019*** | 2.37** |
| Percent Water Area in County Squared | 0.000*** | 0.000 | 0.000 | 0.000** | 0.000*** | 0.000*** | 1.23 |
| Mean Temperature for January 1941-1970 | -0.002 | -0.008 | -0.016*** | -0.013*** | -0.001 | 0.006 | 3.14** |
| Mean Temperature for January 1941-1970 Squared | 0.000* | 0.000 | 0.000*** | 0.000*** | 0.000** | 0.000 | 2.13* |
| Mean Temperature for July 1941-1970 | -0.211*** | -0.044 | 0.153* | -0.004 | -0.125 | -0.207** | 3.57*** |
| Mean Temperature for July 1941-1970 Squared | 0.001*** | 0.000 | -0.001** | 0.000 | 0.001 | 0.001** | 3.54*** |
| Mean Hours of Sunlight in January 1941-1970 | 0.018*** | 0.020*** | 0.023*** | 0.023*** | 0.018*** | 0.017*** | 1.21 |
| Mean Hours of Sunlight in January 1941-1970 Squared | 0.000*** | 0.000*** | 0.000*** | 0.000*** | 0.000*** | 0.000*** | 2.05* |
| Mean Relative Humidity 1941-1970 | 0.007 | 0.038*** | 0.007 | 0.003 | 0.010 | 0.001 | 3.54*** |
| Mean Relative Humidity 1941-1970 Squared | 0.000 | 0.000*** | 0.000 | 0.000 | 0.000 | 0.000 | 3.55*** |
| Plains (=1 if yes) | -0.258*** | -0.151*** | -0.113*** | -0.159*** | -0.304*** | -0.403*** | 9.64*** |
| Tablelands (=1 if yes) | -0.109*** | -0.164** | -0.189*** | -0.077 | 0.034 | 0.040 | 5.19*** |
| Plains with Hills or Mountains(=1 if yes) | 0.017 | 0.041 | 0.037 | 0.000 | 0.047 | 0.032 | 0.62 |
| Hills or Mountains (=1 if yes) | 0.238*** | 0.098** | 0.185*** | 0.184*** | 0.231*** | 0.257*** | 1.92 |
| **County Level Variables** | | | | | | | |
| Mean of County-level Monthly Rainfall | 0.000* | 0.000* | 0.001** | 0.000*** | 0.000 | 0.000 | 1.50 |
| Mean of County-level Monthly Rainfall Squared | 0.000 | 0.000 | 0.000* | 0.000* | 0.000 | 0.000 | 0.32 |
| Standard Deviation of County-level Monthly Rainfall | 0.001*** | 0.001** | 0.001*** | 0.001*** | 0.001*** | 0.002*** | 8.72*** |
| Mean Soil Productivity Index | 0.003*** | -0.002 | 0.002* | 0.003** | 0.004*** | 0.005*** | 3.90*** |
| **Government Payments** | | | | | | | |
| Direct Payments per Owned Acre | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.30 |
| Indirect Payments per Owned Acre | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.65 |
| Disaster Payments per Owned Acre | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.06 |
| Land Retirement Payments per Owned Acre | 0.000*** | 0.000*** | 0.000 | 0.000* | 0.000 | 0.000 | 0.62 |
| Working Land Conservation Payments per Owned Acre | 0.000*** | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 1.01 |
| **Farm Variables** | | | | | | | |
| Value of Production per Operated Acre (Log) | 0.097*** | 0.084*** | 0.086*** | 0.088*** | 0.087*** | 0.079*** | 0.60 |
| Livestock Farms (=1 if yes) | -0.092*** | -0.217*** | -0.158*** | -0.089*** | -0.047** | -0.034 | 8.87*** |
| High Value Crop Farms (=1 if yes) | 0.186*** | -0.010 | 0.070* | 0.165*** | 0.287*** | 0.462*** | 12.51*** |

| Table 3 continued. | | | | | | | |
|---|---|---|---|---|---|---|---|
| Other Variables | | | | | | | |
| Population Interaction Zone in Agriculture | 0.303*** | 0.226*** | 0.228*** | 0.278*** | 0.352*** | 0.431*** | 25.13*** |
| Farms located in Metro County(=1 if yes) | 0.149*** | -0.027 | 0.095*** | 0.130*** | 0.145*** | 0.128*** | 4.81*** |
| Farms located in Rural County(=1 if yes) | -0.176*** | -0.076* | -0.132*** | -0.194*** | -0.221*** | -0.220*** | 3.41*** |
| Year 2006 (=1 if observation in 2006) | 0.053*** | 0.099*** | 0.079*** | 0.073*** | 0.065*** | 0.063** | 0.31 |
| Year 2007 (=1 if observation in 2007) | 0.016 | 0.025 | 0.036* | 0.042** | 0.041* | 0.042* | 0.08 |
| Intercept | 12.722 | 5.511 | -0.937 | 5.667* | 10.417*** | 13.453*** | |
| Regional Dummy Variables (Heartland Region is excluded) | | | | | | | |
| Northern Crescent Region | -0.152*** | -0.411*** | -0.261*** | -0.190*** | -0.156*** | -0.020 | 6.39*** |
| Northern Great Plains Region | -0.925*** | -1.196*** | -1.196*** | -1.160*** | -0.822*** | -0.647*** | 15.11*** |
| Prairie Gateway Region | -0.575*** | -0.688*** | -0.656*** | -0.582*** | -0.449*** | -0.347*** | 5.34*** |
| Eastern Uplands Region | -0.231*** | -0.289*** | -0.405*** | -0.301*** | -0.222*** | -0.013 | 14.52*** |
| Southern Seaboard Region | -0.176*** | -0.295*** | -0.287*** | -0.225*** | -0.118** | 0.005 | 4.78*** |
| Fruitful Rim Region | 0.232*** | 0.053 | 0.083 | 0.294*** | 0.381*** | 0.305*** | 3.36*** |
| Basin and Range Region | -0.023 | -0.543*** | -0.510*** | -0.206** | 0.261** | 0.514*** | 9.85*** |
| Mississippi Portal Region | -0.574*** | -0.456*** | -0.615*** | -0.559*** | -0.570*** | -0.521*** | 2.53** |
| Pseudo $R^2$ ($R^2$ for OLS) | 0.399 | 0.228 | 0.243 | 0.242 | 0.266 | 0.315 | |
| Number of Observations | | | | 15,108 | | | |

***, **, and * indicate statistical significance at 1%, 5%, and 10% levels, respectively.