



AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
<http://ageconsearch.umn.edu>
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

Analyzing the impacts of soil contamination and urban development pressure on farmland values: Unconditional quantile regression estimation

Schreurs, Eloi¹; Peeters, Ludo²; Van Passel, Steven¹

¹ Hasselt University, Faculty of Business Economics, Centre for Environmental Sciences, Martelarenlaan 42, 3500 Hasselt, Belgium

² Hasselt University, Faculty of Business Economics, Econometrics, Research Centre for Entrepreneurship and Innovation, Martelarenlaan 42, 3500 Hasselt, Belgium



**Paper prepared for presentation at the EAAE 2014 Congress
'Agri-Food and Rural Innovations for Healthier Societies'**

August 26 to 29, 2014
Ljubljana, Slovenia

Copyright 2014 by Schreurs, E., Peeters, L and Van Passel, S. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

Abstract

We examine several determinants of farmland values in the Campine region (Belgium) using data on individual sales transactions. Since the study area is well known to have a legacy of heavy-metal pollution of land, one of the focal points is the implicit price of cadmium (Cd) pollution. We use a hedonic pricing model, applying both mean regression (OLS) and quantile regression (QR). Unlike previous hedonic studies using QR, we use unconditional QR (Firpo et al. 2009). The estimates obtained using UQR have a more intuitive interpretation than those obtained using standard (conditional) QR, and allow for less ambiguous comparisons with their OLS counterparts. We find only moderate evidence of price discounts in the middle (median) price range due to Cd contamination. On the other hand, we find that farmland values are strongly affected by the development potential of agricultural land due to urban pressure.

Key Words: Unconditional quantile regression, Hedonic pricing, Farmland values, Soil contamination, Cd pollution, Spatiotemporal price relations, Land development

1. Introduction

In this paper, we estimate major factors influencing farmland prices using a micro-data set covering individual sales transactions realized in a small area within the Campine region (Belgium) over the period 2004–2011. This study area has been chosen for three reasons: (i) the area is known to have a legacy of heavy-metal pollution—in particular, cadmium (Cd) pollution, due to the historic presence of toxic zinc-smelting facilities; (ii) micro-data on individual sales transactions in this area were available from the Belgian Land Registry Office; and (iii) the area is characterized by strong rural-urban interfaces, such that farmland is likely to be highly susceptible to development pressures.

This paper presents an attempt to provide direct evidence of the impact of soil contamination on farmland values. It also addresses the related question of whether agricultural land value is affected by prices of developable (e.g., residential) land. In contrast with earlier hedonic studies using QR, we use recent advances in (non-standard) *unconditional* quantile regressions (UQR), as proposed by Firpo, Fortin, and Lemieux (FFL, 2009). UQR allows researchers to measure the effect of a small change (locational shift) in the covariate X on the quantiles—or any other functional—of the unconditional (marginal) distribution of the dependent variable Y , similar to the way coefficients of a linear model capture the marginal effects on the mean in a standard OLS regression.

Looking at the impact of a particular X on log prices illustrates well the difference between CQR and UQR. Finding that the effect of X estimated using CQR is, say, smaller at the 90th than at the 10th quantile simply means that X reduces the *within-group* price dispersion, where the ‘group’ consists of land plots that share the *same values* of all the covariates other than X —hence, *conditional* effect. This does not mean, however, that increasing X would reduce the *overall* price dispersion as measured by the difference between

the 90th and the 10th quantiles of the *unconditional* price dispersion. To answer this question we have to run a regression of the re-centered influence function (RIF) of the unconditional quantile on the explanatory variables.

Two advantages of UQR using RIF regressions are the following. First, and in contrast with CQR, the marginal effects (implicit prices) can be directly interpreted from the estimation results obtained using UQR-RIF regressions as proposed by Firpo et al. (2009). Second, the latter require only cross-sectional information, whereas previous work applying UQR requires (at least) two points in time and/or the construction of counterfactual distributions (through simulation).

The contribution of this study to the hedonic pricing literature is twofold. First, our study is among the first ones—if not the very first one—dealing with the impact of soil contamination on farmland prices using objective (scientific) measures of pollution. This is an interesting—and thus far largely neglected—issue, considering that—even though it is not visible, does not emit odor or have any other characteristics that easily identify its presence—heavy-metal pollution generates risks to crops grown on the polluted land and, hence, to food safety.

Second, and in contrast with previous empirical studies, we use the *unconditional* (or marginal) quantile regression (UQR) estimation method recently introduced by Firpo, Fortin, and Lemieux (FFL, 2009), thereby taking advantage of the unconditional interpretation it provides.

Apart from these two major contributions, we also examine the land-market segmentation and, thus, the link between prices of agricultural and developable (residential) land; and the importance of spatiotemporal price relations (Maddison, 2009) in the farmland market of the Campine region.

The rest of the paper is organized as follows. Section 2 explains the econometric models, while section 3 gives an overview of the data and the study area. Section 4 presents the results and section 5 concludes.

2 Econometric Model

Linear regression (OLS) has been a standard tool in conducting hedonic pricing analyses for many years. It shows how the expected value—or, more precisely, the conditional mean—of the dependent variable responds to a change in an explanatory variable, other things being equal. However, conventional OLS regression is not well suited to explain the *distribution* of a variable.

Therefore, some researchers have resorted to using standard (conditional) quantile regression (QR) introduced by Koenker and Bassett (1978) to examine how implicit prices vary across the *conditional* distribution of property prices. Specifically, property characteristics may be valued differently at different points of the conditional distribution of the house prices (the so-called quantile effects). Some interesting recent examples (with respect to environmental valuation) are O’Garra and Mourato (2007), Evans and Schaur (2010), and Marques, Fuinhas and Manso (2011).

The linear QR regression model specifies

$$Q_{Y|X}(x, \tau) = x'\beta(\tau)$$

where $Q_{Y|X}(x)$ is the τ -th quantile of the conditional distribution of Y given $X = x$. Consequently,

$$\beta(\tau) = \frac{\partial Q_{Y|X}(x, \tau)}{\partial x}$$

The elements of $\beta(\tau)$ measure the effect of marginally altering the components of x on the τ -th quantile of the conditional distribution of Y on X . In this model, $\beta(\tau)$ is understood as an *unspecified* function of τ (whence its semi-parametric nature). If $\beta(\tau)$ is a positive and monotonically increasing function, this means that increasing X impacts more in higher quantiles of the conditional price distribution. That is, by increasing X , all conditional quantiles move up, but at an increasing rate along the quantiles.

Although CQR has become fairly common, its implications are not always fully recognized, hence easily giving rise to misleading interpretations of the QR results

To address this problem, one should resort to the *unconditional* QR (UQR) method, recently introduced by (Firpo, Fortin and Lemieux, 2009). A gentle introduction and clarifying application can be found in Borah and Basu (2013). This new approach studies the *direct* effect of X on the *unconditional* (marginal) distribution of Y using a RIF-OLS

regression approach. The UQR procedure is designed to answer questions such as ‘what happens to the 90th percentile of the unconditional distribution of Y when X increases?’ Thus, unlike the case of CQR, the answer to this question is *not* conditional on the values of all other covariates.

The RIF-OLS regression approach has some important advantages (Alejo, Gabrielli and Sosa Escudero, 2011): (i) RIF-OLS regressions require less data; that is, only a single cross section is needed—in contrast with previous work requiring repeated cross-section data with at least two time periods and the construction of counterfactual distributions; (ii) RIF-OLS regressions are easier to implement, considering that recovering the unconditional (marginal) price distribution dispenses with the need to use a large number of simulations such as in Machado and Mata (2005), or other numerical solutions such as in Melly (2005); and (iii) the marginal effects can be directly interpreted from the estimation results.

3 Data and Variables

The dataset used to estimate the implicit prices of agricultural land consists of transaction records of individual property sales in 14 municipalities (small urban centers with on average a population size of 20,000) across the Campine region, situated in the north of Belgium just along the Dutch border, that occurred between 2004 and 2011. The dataset, obtained from the Belgian Land Registry Office, is unique and has not previously been made available for research purposes.

Our study area is a peri-urban area, which is known to have a legacy of heavy metal—in particular, cadmium (Cd)—pollution, due to the former presence of toxic zinc-smelting facilities. In addition, rural-urban interfaces are very strong, where ‘urban development’ is replacing agricultural fallows near developed zones. Given the above-mentioned pattern of scattered dwellings (i.e., houses spread out over an extended area) and ribbon development (i.e., houses along both sides of the main provincial roads), every reasonably accessible patch of land is a potential building site.

For this study, 599 observations are used corresponding to transactions that took place from the beginning of April 2004 to the end of December 2011. The transactions were realized through both public auction and private treaty. All observations have been geo-referenced using ArcGIS. The GIS information was used to construct a number of ‘spatial variables’ to be incorporated in the hedonic pricing model, including the housing density within a 1-km radius from each of the plots of land in the sample (to account for urban influences), the spatiotemporal lag of sales prices realized in the recent past, and the distance

of each plot from the Dutch border. The independent variables used in the hedonic pricing model can be subsumed under four headings: (i) transaction-specific characteristics; (ii) neighborhood characteristics; (iii) spatiotemporal lag of sales prices; and (iv) environmental variable. Table 1 lists all the variables included in our hedonic pricing model and provides summary statistics.

The dataset includes nominal sale prices of the farmland, measured in Euros per m². Prices have been converted to real sale prices in constant 2011 Euros (deflated by the general CPI). The size of the farmland plot (acreage) is measured in 1,000 m². Some land transactions have structures (farmhouses, barns, or other agricultural buildings) attached to them, which we indicated by a dummy variable (*structures*=1; vacant=0). Additional indicator variables have been used to typify the farmland plots sold: pasture, arable land, nature zoning, and residential zoning (that is, our sample also includes land that can instantaneously be used for non-agricultural purposes if *residential*=1).

Table 1: Summary statistics

	Variable	Description	Units	Mean	SD	Min.	Median	Max.
Dependent variable	<u><i>price</i></u>	Price farmland	<u>euro per m²</u>	2.82	2.35	0.17	2.34	23.57
Transaction characteristics	<u><i>public</i></u>	Public vs. private sale	=1 if public sale	0.184	0.388	0		1
	<u><i>acreage</i></u>	Farmland plot size	1,000 m ²	14.01	45.52	0.225	6.759	1013
	<u><i>residential</i></u>	Residential zone	=1 if residential	0.078	0.269	0		1
	<u><i>nature</i></u>	Nature(-forest) zone	=1 if nature	0.109	0.311	0		1
	<u><i>pasture</i></u>	Pasture land	=1 if pasture	0.422	0.494	0		1
	<u><i>arable</i></u>	Arable land	=1 if arable	0.262	0.440	0		1
	<u><i>structures</i></u>	Structure vs. vacant	=1 if structure	0.052	0.222	0		1
	<u><i>housedens</i></u>	Housing density ≤ 1 km	1,000 units	0.414	0.340	0	0.326	2.051
Neighborhood (municipality) characteristics	<u><i>farmdensity</i></u>	Farming density	# farms per km ²	1.4	0.9	0.3	1.0	2.7
	<u><i>devlandprice</i></u>	Price of developable land	euro per m ²	97.75	22.02	51.04	98.50	151.13
	<u><i>sptemp_all</i></u>	Sptemp lag 'all' prices	<u>euro per m²</u>	2.84	0.72	1.40	2.72	12.51
	<u><i>sptemp_intra</i></u>	Sptemp lag 'intra' prices	<u>euro per m²</u>	2.84	1.92	0.82	2.45	23.57
	<u><i>distance_NL</i></u>	Distance to Dutch border	<u>km</u>	10.5	6.1	0	9.8	31.1
Environmental variable	<i>Cd</i>	Cd concentration	ppm	1.007	0.688	0.255	0.750	5.436

Since there is no large city/urban center in our study area, we use local housing density as a proxy for urban pressure. This variable is defined as the number of addresses (expressed in 1,000 units) found within a 1-km radius around each farmland plot.

We include two neighborhood characteristics measured at the level of the 14 municipalities of the study area. These municipal-level variables have been primarily included to capture locational influences on farmland values.

The first neighborhood variable is *farming density*, measured as the number of farms per km² in each municipality. The underlying idea is the following. The higher the farming density in a municipality, the better are the local conditions for farming, and, hence, the higher the demand for farmland—hence, potentially giving rise to higher farmland prices in a locality.

The second neighborhood variable is a proxy for *prices of developable land* (to be used for residential, industrial, or other purposes) in each of the municipalities. Here, we basically follow Geniaux, Ay and Napoléone (2011) and assume that agents assess the price of farmland if it becomes developable using location-specific information on developable-land prices. In other words, it is assumed that agents adjust their maximum bid for agricultural land not only depending on its agricultural characteristics, but also—and perhaps primarily—on their anticipation of future land-use zoning. For example, the price of farmland may be influenced by values other than agricultural production, reflecting the fact that farming may not necessarily be the primary motive for owning farmland (Snyder, Kilgore, Hudson and Donnay, 2007). Here, we use the (one-year lagged) five-year moving averages of the (CPI deflated) within-municipal prices per m² of developable land. Given the absence of large city centers in our study area (along with the scattered housing, ribbon development, etc., such that developed space is always nearby and potential land-use spillovers are likely to be very strong), it is not appropriate or feasible to use conventional distance-to-city measures. Second, the intra-municipal prices of developable land could (partly) be reflective of municipal land-use policies. Land markets are subject to competitive pressure as urban centers expand and speculation is frequent. Whether small and marginal farmers could either benefit from these changes or could end up losing access to land, depending on the land-rights system.

As a final transaction characteristic, we include the (inverse of) *distance* of the farmland plot to the Dutch border. The reason for including this distance measure is the existence of a ‘border effect’ (cross-border sales), caused by the fact that over the period 1995–2010 farmland prices in the Netherlands have been significantly higher than in Belgium (on average about 60% more expensive). Therefore, many Dutch farmers living close to the border would want to buy farmland just across the border in Belgium.

We also include a *spatiotemporal lag* of farmland prices, basically following Maddison (2009). This variable is defined as a spatially weighted average of realized farmland prices in the recent past within some predefined neighborhood.

Finally, we include a measure of Cd contamination of land by using results from an earlier

trend-surface analysis (Schreurs, Voets and Thewys, 2011). The trend surface was estimated through ordinary *kriging*, on the basis of about 12,000 soil samples.

The metallurgic industry in the area has diffusely contaminated the Campine region in Belgium with various heavy metals, particularly cadmium (Cd). The cadmium (Cd) concentration of farmland (measured in ppm) is the sole environmental variable included in our analysis. We enforce the impact of Cd contamination to be *identically* equal to zero at Cd levels below 2 ppm because the legal threshold value for Cd contamination in agricultural soils is 2 ppm. Therefore we use a threshold regression, given (simplified) by

$$\ln p_i = \alpha + \beta c_i I(Cd_i > \theta) + \text{other covariates} + u_i$$

where p_i is the price per m^2 of farmland, and the indicator function $I(\bullet)$ is given by

$$I(Cd_i > \theta) = \begin{cases} 0 & \text{if } Cd_i \leq \theta \\ 1 & \text{if } Cd_i > \theta \end{cases}$$

where $c_i = [\ln(Cd_i/2)]^2$, and θ is the exogenously set critical threshold value of 2 ppm. Thus, the observations are divided into two ‘regimes’ depending on whether the level of Cd is smaller or larger than the threshold value θ , where the regimes are distinguished by differing slopes; that is, $\beta = 0$ (no effect) for $Cd_i \leq \theta$ and $\beta < 0$ (negative effect) for $Cd_i > \theta$.

4 Results

In this section, we present the estimation results. A summary of the results is provided in Table 2. Throughout the analysis, we have consistently used a log-linear specification of the empirical model. Since the dependent variable is measured in logs, we look at proportionate (or percentage) changes in the sales price per square meter. A standard Moran’s I test indicated the presence of spatial error autocorrelation for OLS and CQR (not possible, however, for UQR) when the spatiotemporal lag was excluded. The error autocorrelation was lifted after including the spatiotemporal lag.

The OLS results are presented in column 0 of Table 2. They are reported for comparative purposes only because conclusions based on the OLS estimates are vastly oversimplified; that is, heterogeneous buyers may value the farmland plot characteristics and/or neighborhood attributes differently. Evidently, conventional mean regression is not capable of detecting heterogeneity along the price distribution.

The CQR results are reported in the odd-numbered columns 1–9 of Table 2, while the UQR results are given in the even-numbered columns 2–10 of Table 2. In what follows, we are primarily interested in the estimates obtained using UQR.

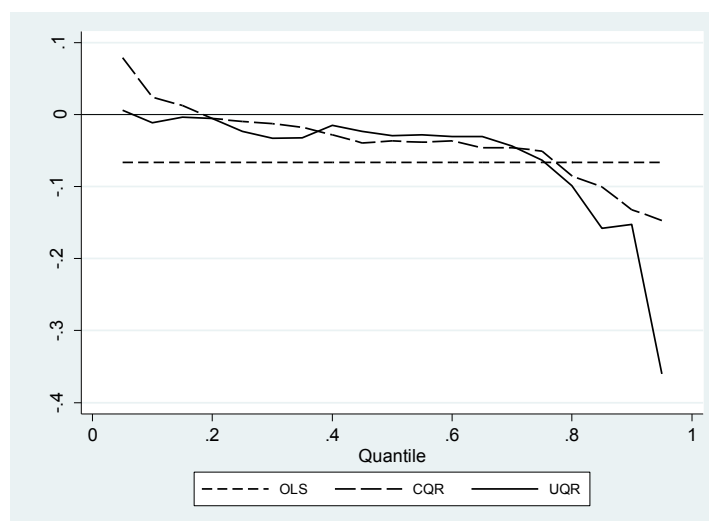
We focus on a key set of explanatory variables that enable us to identify the unconditional effects (implicit prices) of urban development pressure to and Cd pollution of the soil. The other variables will only be discussed in passing (only an abbreviated discussion).

To begin with, we compare the estimation results from mean regression (OLS) and quantile regressions (CQR and UQR). Therefore, we look at the estimates for the impact of plot size on farmland values. In panel A of 1, we present the OLS estimate (dashed horizontal line) and Q-plots, which show the CQR quantile effects of (log) plot size at selected points of the conditional distribution of the (log) sales prices (long-dashed curves). The OLS estimate predicts a uniform *locational shift* downward, which represents a constant elasticity estimate of -0.066 . On the other hand, the quantile effects returned by CQR are monotonically decreasing across the *conditional* distribution of farmland prices given plot size; that is, a larger plot size has a stronger negative effect on the value of *conditionally* high-priced farmland plots (-0.132 at the 90th percentile) than on the value of *conditionally* low- and medium-priced farmland plots (-0.036 and zero at the 50th and 10th percentile, respectively)

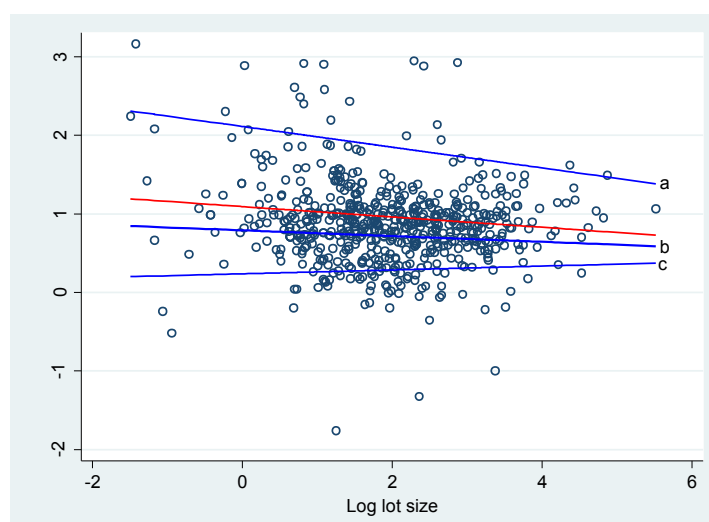
While the presentation of Q-plots is quite common in the empirical QR literature, their implications are not always fully recognized—while mostly being misinterpreted. Therefore, in order to help the interpretation of the CQR, panel B of Figure 1 shows the OLS (conditional mean red regression line) and CQR (conditional percentile blue regression lines). The slopes of the straight lines in panel B of Figure 1 indicate how the value of Y changes with X (i.e., at each value of X) as one moves along the corresponding percentiles of the *conditional* distribution. The lines are converging; that is, the conditional distribution of Y is less spread out at higher values of X than at lower values of X . The negative effect of plot size on sales prices tends to be larger (in absolute value) for higher-prices farmland plots than for lower-priced plots; that is, the quantile regression for the upper part of the conditional price distribution (90th percentile) reveals a significant negative effect of plot size on sales prices, whereas the quantile regressions for lower quantiles of the conditional price distribution (50th and 10th percentile) reveal little or no effect. Note that the conclusion from OLS (red line) that *all* prices would fall (shift downward) with increasing plot size turns out to be a vast oversimplification.

Figure 1: Regression results for impact of plot size on sales prices

A: Q-plot (quantile effects)



B: OLS mean and CQR percentile regression lines



Notes: The red line represents the predicted conditional mean obtained using OLS. The blue lines represent the predicted 10th, 50th, and 90th percentiles of the conditional distribution obtained using CQR. The predictions are conditioned on the mean value of all other covariates (at the raw/un-logged scale), which differ from the effects on the unconditional quantiles.

Point $a = 1.383$ on the 90th percentile line of the *conditional* distribution of sales prices corresponds to the 86th (almost the 90th) percentile of the *unconditional* distribution of sales price, while point $b = 0.589$ on the 50th percentile line corresponds to the 24th percentile of the *unconditional* distribution of sales prices. Thus, point b is located in the medium-price range of the conditional distribution of log sales prices, *given* a log plot size equal to 0.589 (which is about 1,800 m²), whereas it is located in the lower-price range of the unconditional

distribution of log sales prices. This finding obscures the interpretation of the CQR results, and makes it difficult to see how changes in the distribution of (log) plot sizes translate into the unconditional distribution of (log) sales prices.

The UQR results, obtained through the use of RIF-OLS, show the effects on the *unconditional* (or marginal) distribution of (log) sales prices. It can be seen that the UQR quantile effects are also monotonically decreasing across the quantiles of the unconditional distribution, but the decline is now more pronounced, particularly in the upper-price range. The quantile effect goes from about zero for the lowest-priced farmland plots to -0.152 at the 90th percentile of the unconditional distribution. The UQR results are more directly interpretable, because now they properly suggest that the negative effect of plot size is stronger for high-priced farmland plots. A possible explanation (interpretation) of those results is that certain buyers (among a heterogeneous group of buyers), who are prepared to pay a high price per m^2 , are willing to bid prices further up in order to get a smaller piece of farmland, possibly because of *speculative* motives—in anticipation of future land-use conversions allowing for urban development. Moreover, it has been found in earlier empirical work (e.g., Cavailhès and Wavresky (2003)) that the *potential* for urban development increases with decreasing plot size.

Furthermore, we found similar effects with other variables that are closely related to (potential) urban development. The implicit prices of farmland plots located in residential-zoning areas are strongly influenced only in the upper range of the unconditional price distribution, which is not really surprising. The quantile effects of housing density within a 1-km radius around farmland plots, is monotonically increasing along the unconditional price distribution, whereas the quantile effects of the (intra-municipal) prices of developable land, are monotonically decreasing. The effect of distance to the Dutch border is much stronger in the upper range of the unconditional price distribution. Finally, spatiotemporal effects turn out to be significantly affecting only medium-priced farmland plots.

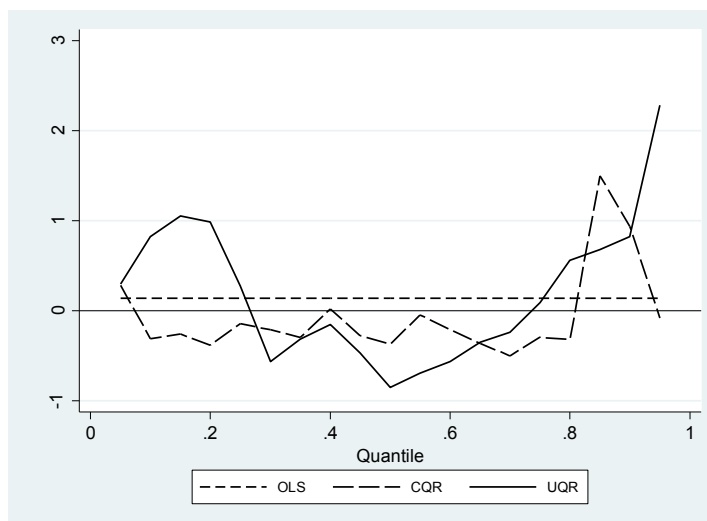
Next, we look at the impact of Cd contamination on sales prices. The OLS estimate predicts a uniform upward (but insignificant) shift, which is a hardly credible result. Conversely, as shown in panel A of Figure 2, both the CQR and UQR results reveal a U-shape pattern of the estimated quantile effects, although this pattern is more pronounced for UQR. Curiously, the effect of Cd contamination is negative and significant only in the case of medium-priced farmland plots at a Cd-concentration level greater than 2 ppm (recall that the effect was set to zero at levels smaller or equal than 2 ppm, by construction). The UQR

estimate at the 50th percentile (median) of the unconditional price distribution suggests an elasticity of the sales price with respect to Cd contamination of -0.694 at a Cd level of 3 ppm, and an elasticity of -1.186 at a Cd level of 4 ppm. Recall that the elasticity is calculated as $2\hat{\beta}_{Cd}(q_{\tau} = 0.5) \times \ln(Cd/2)$. Panel B of Figure 2 shows the OLS (conditional mean) regression line, along with the CQR (conditional percentile) regression lines. The graph suggests a widening of the conditional price distribution, though it should be remembered that the slopes of both the 10th and 90th percentile curve are not statistically significant. On the other hand, the results suggest that the unconditional distribution of sales prices becomes more positively skewed at higher levels of Cd contamination. This finding suggests that remediation of the soil (i.e., reduction of Cd level) in the study area would render the unconditional distribution of log prices more symmetric. In summary, there is only weak empirical evidence of negative effect of Cd contamination. Moreover, the counter-intuitive (positive) signs of the UQR estimates at the lowest range (significant) and at the top range (not significant) of the unconditional price distribution warrant further research.

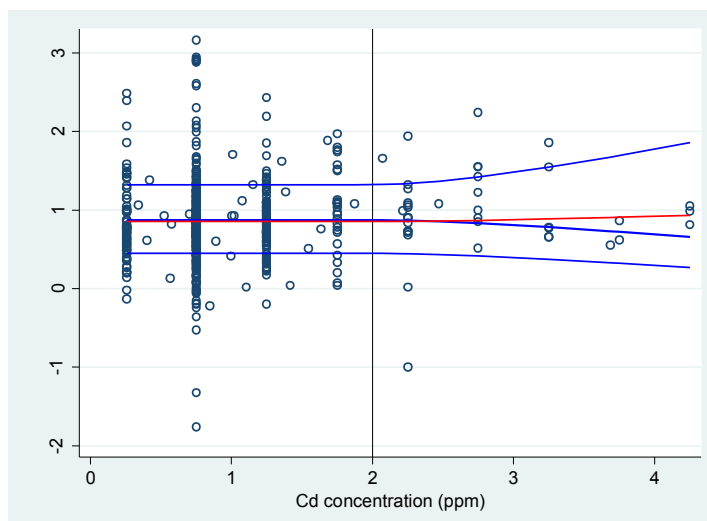
Our results concerning the impact of Cd contamination on agricultural land prices are mixed. This could be explained by the fact that certain farmers (e.g., dairy farmers) might not take the soil contamination into account because they use the land for manure disposal. Another possible explanation is that individuals have bought these pieces of farmland with non-agricultural interests (e.g., horses), although all observations are labeled as agricultural land. A last explanation can be that farmland buyers might have been unaware of the exact heavy metal concentrations present in the acquired parcels and its resulting land use restrictions.

Figure 2: Regression results for impact of Cd contamination on sales prices

A: Q-plot (quantile effects)



B: OLS mean and CQR percentile regression lines



Notes: The red line represents the predicted conditional mean obtained using OLS. The blue lines represent the predicted 10th, 50th, and 90th percentiles of the conditional distribution obtained using CQR. The predictions are conditioned on the mean value of all other covariates (at the raw/un-logged scale), which differ from the effects on the unconditional quantiles.

5 Conclusions

In this paper, our aim was to go beyond conventional OLS estimation, and use quantile regression (QR) to estimate major determinants of farmland values in the Campine region, situated in the north of Belgium. The QR framework provides a pragmatic approach to examine the impacts of transaction characteristics along the entire price distribution. Our specific contributions were (i) the focus on soil pollution, and (ii) the use of (non-standard) unconditional QR (UQR) proposed by Firpo, Fortin and Lemieux (2009). The UQR approach is better suited to address policy issues and welfare implications of, for example, soil remediation.

The fact that our study area is known to have a legacy of Cd pollution of the soil raises a concern about welfare losses associated with lower farmland values. It was found that Cd contamination has different effects at different points of the farmland-price distribution, with a significantly negative impact on the median price (or the central region of the price distribution). However, we were able to find only weak empirical evidence in support of a negative effect on prices of farmland. On the other hand, we found compelling evidence of increasing farmland values due to urban development forces. Interestingly, farmland plots in the high-price range of the unconditional distribution are positively influenced by densely populated surroundings, while prices of developable land spill over to farmland plots in the low-price range of the unconditional distribution.

We also showed that conventional OLS leads to conclusions that are vastly oversimplified. We applied a simple method to estimate the effect of changes in the distribution of explanatory variables on unconditional quantiles of the distribution of the outcome variable. In its simplest version, the method consists of running a regression of the re-centered influence function of quantiles on the explanatory variables, to so-called RIF regression. The RIF regression enabled us to easily recover the marginal effects of changes in the covariates distribution on the unconditional quantiles of the sales prices (Firpo, Fortin and Lemieux, 2009). Therefore, we recommend using the UQR approach, especially given the ease of interpretation of the results.

Table 2: Estimation results from mean regression (OLS) and quantile regressions (CQR and UQR) – Dependent variable: Log sales price per square meter

	OLS	10th Quantile		25th Quantile		50th Quantile (Median)		75th Quantile		90th Quantile	
		CQR	UQR	CQR	UQR	CQR	UQR	CQR	UQR	CQR	UQR
	(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<i>Transaction (plot-level) characteristics</i>											
Public sale	0.217*** (0.063)	0.086 (0.124)	0.121 (0.087)	0.180*** (0.068)	0.148** (0.061)	0.254*** (0.055)	0.186*** (0.048)	0.259*** (0.074)	0.441*** (0.076)	0.324** (0.129)	0.143 (0.124)
Log plot size	-0.066*** (0.025)	0.025 (0.047)	-0.011 (0.033)	-0.009 (0.025)	-0.023 (0.025)	-0.036** (0.017)	-0.029 (0.019)	-0.051 (0.032)	-0.063** (0.027)	-0.132*** (0.045)	-0.152*** (0.048)
Structures	0.162* (0.093)	0.094 (0.256)	0.152 (0.117)	0.197 (0.146)	0.137 (0.092)	0.248** (0.097)	0.172** (0.087)	0.203* (0.104)	0.242* (0.135)	0.218* (0.116)	0.093 (0.220)
Pasture land	-0.157*** (0.048)	-0.225** (0.100)	-0.171** (0.087)	-0.215*** (0.059)	-0.190*** (0.063)	-0.114*** (0.031)	-0.151*** (0.045)	-0.066 (0.049)	-0.114** (0.058)	-0.121 (0.089)	-0.023 (0.094)
Arable land	-0.058 (0.056)	-0.055 (0.115)	-0.090 (0.104)	-0.103 (0.066)	-0.102 (0.068)	-0.089** (0.042)	-0.084* (0.051)	-0.021 (0.063)	-0.041 (0.071)	-0.070 (0.142)	-0.049 (0.115)
Nature-forest zone	-0.100 (0.075)	-0.288* (0.153)	-0.218 (0.143)	-0.176** (0.078)	-0.156* (0.089)	-0.146* (0.076)	-0.116* (0.062)	-0.097 (0.069)	-0.116* (0.070)	-0.084 (0.220)	0.039 (0.131)
Residential zone	0.280** (0.111)	0.055 (0.156)	0.012 (0.145)	0.013 (0.117)	0.135 (0.090)	0.182 (0.120)	0.057 (0.077)	0.296 (0.223)	0.171 (0.113)	0.600** (0.274)	0.605*** (0.231)
Housing density ≤ 1 km	0.243*** (0.075)	-0.005 (0.179)	-0.012 (0.105)	0.146* (0.087)	0.067 (0.077)	0.148** (0.060)	0.169*** (0.060)	0.295** (0.128)	0.289*** (0.092)	0.426** (0.174)	0.470*** (0.180)
<i>Neighborhood (municipality-level) characteristics</i>											
Log number of farms per km ²	0.071** (0.035)	0.093 (0.078)	0.117* (0.061)	0.099** (0.039)	0.096** (0.041)	0.091*** (0.028)	0.097*** (0.030)	0.048 (0.034)	0.045 (0.039)	0.105 (0.067)	-0.047 (0.067)
Log price per m ² of developable land	0.275** (0.137)	0.625*** (0.189)	0.646*** (0.220)	0.596*** (0.118)	0.637*** (0.141)	0.479*** (0.124)	0.500*** (0.101)	0.227 (0.163)	0.085 (0.145)	0.152 (0.241)	-0.095 (0.261)
Log spatiotemporal lag of sale prices per m ²	0.134* (0.081)	0.243* (0.146)	0.059 (0.112)	0.174** (0.067)	0.190*** (0.068)	0.108 (0.069)	0.109** (0.050)	0.146 (0.096)	0.147* (0.075)	0.243** (0.111)	0.099 (0.145)
Inverse distance to Dutch border	0.116*** (0.032)	0.072 (0.063)	0.027 (0.032)	0.063 (0.049)	0.061** (0.025)	0.071 (0.047)	0.061** (0.030)	0.128** (0.051)	0.168*** (0.061)	0.198*** (0.074)	0.423*** (0.133)
<i>Environmental (plot-level) variable</i>											
[Log (Cd/2)] ₂ ² for 2 < Cd ≤ 4.25	0.138 (0.438)	-0.310 (0.503)	0.823** (0.317)	-0.145 (0.537)	0.274 (0.669)	-0.373 (0.798)	-0.856** (0.390)	-0.295 (1.300)	0.091 (0.621)	0.937 (1.168)	0.822 (1.327)
Number of observations	511	511	511	511	511	511	511	511	511	511	511
Adjusted R ²	0.172		0.028		0.094		0.141		0.175		0.139
Pseudo R ²		0.098		0.124		0.124		0.152		0.207	

Notes: Robust standard errors for OLS and UQR (RIF-OLS) and bootstrapped standard errors for CQR (based on 2,000 bootstrap replications) are given in parentheses. The estimated constant terms are not reported in the table to save space. Outliers, for which Cd > 4.25 ppm, have been discarded. * Significant at 10% level; ** Significant at 5% level; *** Significant at 1% level.

References

- Maddison, D. (2009). A Spatio-temporal Model of Farmland Values. *Journal of Agricultural Economics* 60: 171-189.
- Koenker, R. and Bassett, G. (1978). REGRESSION QUANTILES. *Econometrica* 46: 33-50.
- O'Garra, T. and Mourato, S. (2007). Public Preferences for Hydrogen Buses: Comparing Interval Data, OLS and Quantile Regression Approaches. *Environmental and Resource Economics* 36: 389-411.
- Evans, M. F. and Schaur, G. (2010). A quantile estimation approach to identify income and age variation in the value of a statistical life. *Journal of Environmental Economics and Management* 59: 260-270.
- Marques, A., Fuinhas, J., Manso, J. (2011). A Quantile Approach to Identify Factors Promoting Renewable Energy in European Countries. *Environmental and Resource Economics* 49: 351-366.
- Firpo, S., Fortin, N. M., Lemieux, T. (2009). Unconditional Quantile Regressions. *Econometrica* 77: 953-973.
- Borah, B. J. and Basu, A. (2013). HIGHLIGHTING DIFFERENCES BETWEEN CONDITIONAL AND UNCONDITIONAL QUANTILE REGRESSION APPROACHES THROUGH AN APPLICATION TO ASSESS MEDICATION ADHERENCE. *Health Economics* 22: 1052-1070.
- Alejo, J., Gabrielli, F., Sosa Escudero, W. (2011). The distributive effects of education: an unconditional quantile regression approach. *MPRA Paper No. 42933 available at <http://mpra.ub.uni-muenchen.de/42933/>*.
- Machado, J. A. F. and Mata, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics* 20: 445-465.
- Melly, B. (2005). Decomposition of differences in distribution using quantile regression. *Labour Economics* 12: 577-590.
- Geniaux, G., Ay, J.-S., Napoléone, C. (2011). A SPATIAL HEDONIC APPROACH ON LAND USE CHANGE ANTICIPATIONS*. *Journal of Regional Science* 51: 967-986.
- Snyder, S. A., Kilgore, M. A., Hudson, R., Donnay, J. (2007). Determinants of Forest Land Prices in Northern Minnesota: A Hedonic Pricing Approach. *Forest Science* 53: 25-36.
- Schreurs, E., Voets, T., Thewys, T. (2011). GIS-based assessment of the biomass potential from phytoremediation of contaminated agricultural land in the Campine region in Belgium. *Biomass & Bioenergy* 35: 4469-4480.
- Cavailhès, J. and Wavresky, P. (2003). Urban influences on periurban farmland prices. *European Review of Agricultural Economics* 30: 333-357.