

Determination and forecast of agricultural land prices

Bastiaan Philip Reydon*

Ludwig Einstein Agurto Plata**

Gerd Sparovek***

Rafael Guilherme Burstein Goldszmidt****

Tiago Santos Telles*****

Keywords

land market, hedonic prices, multiple regression analysis, land policy

JEL Classification

R14, R15

Palavras-chave

mercado de terras, preços hedônicos, análise de regressão múltipla, política fundiária

Classificação JEL

R14, R15

Abstract

The aim of this study is to discuss and apply hedonic methodology for the determination and forecast of land prices in specific markets. This is important due to the fact that there is no official or reliable information in Brazil on current prices in land market transactions. This hedonic price methodology uses a multiple regression model which has, as an explanatory variable, the price per hectare and independent variables related to physical attributes (soil, climate and terrain), production (systems of production, location, access), infrastructure of the property and expectations (regional scenario, local investments). Application of the methodology to a Homogeneous Zone of the state of Maranhão, in Brazil, generated a parsimonious model, in which five independent variables were responsible for 70% of the variance in the price of agricultural land.

Resumo

O objetivo do presente trabalho foi discutir e aplicar a metodologia de preços hedônicos para determinação e previsão do preço da terra rural em mercados específicos. A sua importância decorre do fato de que não existem no Brasil informações oficiais ou fidedignas sobre preços praticados nos negócios com imóveis rurais. Essa metodologia de preços hedônicos utiliza um modelo econométrico de regressão múltipla, tendo como variável dependente o preço por hectare, e como variáveis explicativas as relacionadas ao meio físico (solo, clima e relevo), à produção (sistemas de produção, localização, acesso), à infraestrutura do imóvel e às expectativas (situação regional, investimentos locais). A aplicação de tal metodologia a uma zona homogênea do Estado do Maranhão implicou um modelo parcimonioso, em que cinco variáveis independentes permitiram a explicação de 70% da variância do preço por hectare da terra rural.

.....
*Professor Livre Docente IE/UNICAMP

**Professor Pleno FATEC

***Professor Titular ESALQ/USP

****Professor Adjunto EBAPE/FGV

*****Pesquisador IAPAR

1 Introduction

In Brazil, prior to the economic stabilization of 1995, the issue of determination of agricultural land prices was frequently relegated to the sidelines for two reasons: firstly, it was considered to be a preoccupation only of landowners and secondly, due to the difficulties in calculation on account of the high inflation. However, the result of Brazilian land policies and disputes in the courts, amongst other issues, have been demonstrating the importance that the determination and forecasting of adequate rural and urban land prices exerts in Brazil today.

Land market dynamics and the consequences of the evolution of its prices have played a crucial role in the aims and goals of land policies and land administration. For instance, the sharp reduction in agricultural land prices, following the *Plano Real* in 1995, significantly favored the attainment of the goals for land reform in the first term of President Fernando Henrique Cardoso (Delgado, 2005; Sallum Junior, 1999). Along with the democratization of land through the market, which began in 1997, the programs *Cédula da Terra* (Land Bill), *Banco da Terra* (Land Bank) and *Crédito Fundiário* (Land Credit) had, because of acquaintance with the markets and construction of suitable policies, many different impacts. While the *Cédula da Terra*, because of subsidies offered with these agreements made at lower than market price, reduced market prices, the Banco da Terra, without subsidies and without price supervision, expanded the demand for land, thus significantly increasing its price (Reydon and Plata, 2006).

In various legal skirmishes, the expropriation of land for land reform and land purchases for the programs called “access to land through the market” have been concluded using market prices. But how is market price

established? How to unravel the variables that determine the dynamics of the land price in a specific geographical location or local market? What kind of model should be used to forecast this price? This article contributes with a methodology for forecasting the land price in specific markets with the aim of providing supporting data, amongst other aims, to the agricultural policymakers in charge of democratizing access to rural land in Brazil.

In the specialized international literature on agricultural economics, empirical papers like those of Peters (1966), Lloyd, Rayner and Orme (1991), Lloyd (1994) and Hallan, Machado and Rapsomanikis (1992), concentrate their explanation of agricultural land price dynamics from a macroeconomic perspective. These authors recognize that agricultural land is an asset and that its price is determined by the capitalization of future income obtained from its productive and speculative use. Productive incomes are derived from agricultural products while speculative incomes derive from their characteristic as an asset that maintains value over time. In the case of Brazil, because of the experience of high inflation, speculative use has had a large impact as shown by studies like those of Pinheiro (1980), Reydon (1984), Brandão (1986) Brandão and Rezende (1989), Bacha (1989), Reydon (1992), Reydon and Romeiro (1994) and Plata (2001).

Based on these studies, this article develops a methodology for determining land prices for specific markets, taking into account both of the aforementioned characteristics of land use. It starts by selecting the determinants of land prices in specific markets and then establishes econometric models to explain the dynamics and forecast price of land in homogenous areas¹.

2. Theoretical parameters

2.1 Hedonic price

Hedonic price analysis is a statistical technique developed more than seventy years ago to assess product quality issues. There are two basic approaches in the literature to understanding price characteristics.

One tradition relates this price to a consumer's willingness to pay for a characteristic. This utility-based interpretation is reflected in the use of the term hedonic to describe the approach, and was the original view of the subject matter adopted by Andrew Court (Goodman, 1998) and other early practitioners. Lancaster (1966) proposed a theory of consumer utility based on characteristics rather than on goods, thus it is possible to establish a relationship between the value of the goods and their characteristics. From this, Lancaster proposed the existence of two stages in the relationship between individuals, goods and their characteristics: one between goods and their characteristics (technical relationship) and the other between individuals and the characteristics of the goods (individual preference relationship). Pendleton and Mendelsohn (2000) described the rather restrictive conditions under which the hedonic function can be derived from an underlying utility function.

On the other hand, the second approach, developed by Rosen (1974), has generally been accepted as the paradigm of the hedonic approach. Rosen relates the hedonic function to the supply and demand for the individual characteristics of each commodity. The hedonic approach is a method that estimates a function that relates the price of the commodities to the different attributes that it possesses (implicit price).

Rosen (1974) based his argument on two pillars: firstly the fact that the product has a price and secondly that it has

measurable characteristics or attributes which define the so-called hedonic price or implicit price. Rosen (1974) also assumes that consumers purchase one single unit of the asset with its particular characteristics. As the author states, additions in income always increase maximum utility and therefore it should be expected that consumers with higher incomes will purchase larger quantities of characteristics. However, in general, there is no reason why the quantities required for all features must always increase with income, since some of their components may increase and others decrease. Consequently therefore, the model has a natural market segmentation, where consumers acquire cash to buy similar products with similar characteristics.

According to this theory, a class of goods that are described by their n attributes or characteristics defines the competitive market. The components of this vector are thus measured as each consumer assesses each characteristic equally. However, there are differences in the valuation of each 'features package' for each agent market. Each product has a share of the market price and is associated with a fixed value of the vector q , revealing an implicit function $p(q) = p(q_1, q_2, \dots, q_n)$ relating prices and characteristics. This function is equivalent to hedonic regression rates obtained by comparing search rates with different characteristics.

Rosen's approach is similar to the one with competitive market segmentation, using data values to find products and their characteristics, based on a spatial equilibrium model. Some research procedures, based on the concept of spatial heterogeneity, recommend the creation of homogeneous submarkets for hedonic price models (Abraham et al., 1994; Bowen et al., 2001).

This view was explored further by many authors, including Triplett (1983), Epple (1987), Feenstra (1995) and Pakes (2002). The methodology has recently been used extensively in real estate, with Plantinga and Miller (2001), Bastian et al. (2002), Angelo et al. (2004), Taylor and Brester (2005), Arraes and Souza Filho (2008), Guiling et al. (2009), Kostov (2009), Sander and Polasky (2009), Deaton and Vyn (2010), Ma and Swinton (2011) and Jaeger et al. (2012).

2.2_Agricultural land price determination

The price of agricultural land, in a specific geographical area, reflects the existing market structure and the political and socioeconomic development of the region. Market prices guide the private economic agents in the land market in purchases and sales; they are also a reference for the government in its rural democratization of access to land and in land taxation programs; they are used by credit institutions for the computation of the mortgage and land valuation as a guarantee for rural loans. Consequently the price of land is, on the one hand, the relevant variable that expresses the expectations of the economic agents for this resource and, on the other hand, it acts as a signal to be considered by the policy makers when it is proposed to define efficient economic and social land use and distribution.

But how to estimate and describe the dynamics of prices in imperfect land markets, as is the case in Brazil, in which land has a fixed, immovable and concentrated supply? On the one hand, land can be used as a productive factor in the production of rural goods and, on the other, as a speculative asset, as it maintains value from one period to another. There are also rules concerning its usage (for instance, the legal forest reserve) and taxes on properties, besides the cultural and socio-political characteristics that affect the market. In

this context, the rural land price synthesizes the effect of all the factors that interact in its market. Therefore, this paper discusses, in theoretical and empirical ways, the determinant variables for land prices and the dynamics of the land market in Brazil.

Theoretically, it is assumed that these land markets are established in capitalist economies in which the economic agents have expectations and make decisions to obtain maximum monetary gain². In this scenario, of enterprise and market economies, the owners of wealth obtain different kinds of assets³, with different levels of liquidity to obtain monetary gains and protection from the uncertainties of the capitalist economy and try to predict the psychology of the markets and decide whether or not to buy the assets that, according to their expectations, will provide higher net returns (Reydon, 1992; Plata, 2001; Reydon and Plata, 2006).

Rural land is an important asset because it possesses three particular characteristics: (i) scarcity; (ii) physically immobility; and (iii) durability (Dasso et al., 1995). The scarcity of land is not only a consequence of its physical scarceness, but also the scarcity of the products that emanate from it. However, being an immobile factor that cannot be reproduced, the economic scarcity of land is caused by its low elasticity of production and substitution, which can be privately appropriated by some agents. Nevertheless, the development of technologies that increase its productivity, as well as administrative measures such as land reform, for example, can substantially modify the level of land scarcity in a region (Plata, 2006).

It is also assumed that a land market is created when the ownership of the region is accepted in general terms, regardless of the way it is maintained or the guaranties for its maintenance⁴ (Binswanger, 1993). Therefore,

with any changes in legislation or in the guarantees that a property may have⁵, its condition as an asset becomes more uncertain, increasing the risk associated with acquisition and decreasing the liquidity, rate of capitalization and its price (Deininger and Feder, 2001). The reference assumed here has always been the property, irrespective of its form, because in some areas or countries, where the property is not formally established but socially accepted and land is traded, there is a land market (Binswanger et al., 1995).

Land prices are the result of a trade between purchasers and sellers in the land markets, but this trade only occurs when a purchaser has higher expectations than the seller about the future gains from that land. Consequently, the changing expectations of future gains from the land and, therefore, its price, are the most important variables in understanding the dynamics of the land market (Case and Quigley, 1991).

In summary, rural land can be characterized as being, simultaneously, a capital and a liquid asset, negotiated at flexible prices – established by the capacity of the owners to accumulate the asset. The main reason for this is that the supply of land is fixed⁶ and the market price will be determined by the dynamics of demand.

The expectations of the owners can determine the quantity of land to be negotiated, but the purchasers' expectations of future gains with the use of the land is what will establish the price. In this context, according to Reydon (1992), similar to all assets, the price of rural land is an expression of the prospective gains for the three capitalized attributes:

$$P = q - c + l \quad (1)$$

Where:

- q – productive quasi-rents: the expected gains from productive uses of the property. The value of this attribute depends on the expected gains from rural production and the possibility of other gains resulting from possession of the land, such as credits or government subsidies.
- c – maintenance costs: expected costs of maintaining the land in the portfolio of the agent; this means all the non-productive costs associated with the property, such as the transaction costs, land taxes and the like.
- l – liquidity premium: the ability to sell the land in the future. This is the least objective part of the price computation and is primarily formed by the agents' expectations in relation to the land markets. It is higher when the economy grows and the demand for land as a capital asset increases, or when there is an increase in the demand for liquid assets. Sometimes, in a crisis, when expectations for other liquid assets are worse than they are for land, its liquidity may also grow.

It is important to emphasize that the specific, local Brazilian land markets are imperfect mainly because: a) of a significant political and social inequality of property distribution; b) an individual economic agent can manipulate the supply and the price of land; c) the landless need land but they are economically unable to obtain it; d) land is not a simple product, the properties have different dimensions, quality, fertility and surfaces; e) there are spatial conditions that affect the price (Plata, 2006; Reydon, 2011). Empirical evidence shows, however, that regions with dynamic land markets also have dynamic product, labor and credit markets (Case and Quigley, 1991; Alston et al., 1996; Lambin et al., 2003; Barbieri and Bilborrow, 2009).

It is important to emphasize that land markets have two different segments: the trade market and the rental market. On the one hand, an economic agent that operates in the trade market is willing to pay for the total possible gains: the productive quasi-rents and the liquidity premium of the land. On the other hand, renters will be willing to pay a rent based just on productive profit and, because of this, the value of the rent of the land can be considered as a proxy variable of its productive gains.

2.3 Variables in land price determination

Based on the aforementioned theory, it can be stated that the land price in a specific market is determined by the expected productive and speculative gains from the property. The main variables that explain the dynamics of these gains and the land prices are:

- The overall demand and prices for products from specific farming activities. This demand is determined by prices of products and by input costs such as: technology, mechanization (capital) and other factors used in production. In microeconomic terms, the productive profit from land use at a particular moment in time would be similar to the expected value of the land's marginal product⁷. So, the productive gain from land would depend on the market conditions for the product and the technical conditions for production, because the land's marginal physical productivity is a consequence of a technical relationship with other factors in a specific technology. An increase in the price of the product, due to an increase in profit or a change in consumer preferences, creates expectations of an increase in productive profit.

The same occurs when production costs fall (in the case, for instance, of a decrease in the price of assets, ease of access to capital, improvement in technology and/or in the conditions of production), which increases the production function and the physical productivity of the land.

- The large increase in land use for food and energy production around the world in the last ten years has had a big impact on demand and the price of land in all developing countries, particularly in Latin America and Africa (Cotula, 2008 and 2011; Msangi and Ewing, 2008; Deininger, 2011).
- The infrastructure of production and trade affects the expected productive gains from land. The existence of irrigation infrastructure, availability of water, access, transportation, proximity to the centers of consumption and information has a positive effect on land prices, as well as decreasing the risks to its productive gains. In many cases, these variables determine the different land prices locally.
- Institutional restrictions on the utilization of land create negative expectations about productive gains, decreasing the price of the land. Good examples are the Laws of Forestry Preservation (Forest Code) that reduce land prices. On the other hand, the social benefits from the preservation of the environment can be high and the alternative use of rural land, such as ecological tourism, can generate optimistic expectations of increased gains from land.
- Another variable that affects the land price is the level of fragmentation; the smaller the properties the higher the prices, mainly because of the increased liquidity. In the case of agricultural

land, the impact of fragmentation on land prices depends on the area required for efficient agricultural exploration in the region (Reydon et al., 2006).

- Population growth can have an important effect on land prices for at least two different reasons: an increase in demand for farming products (food) and space for urbanization and leisure. The increase in demand for land for non-farming purposes mostly increases prices only with a Homogeneous Zone.
- Inflation affects land prices in two ways: firstly, by changing productive gains, due to the increase in the price of products and inputs⁸. The second and more important way relates to land's capacity to retain value derived from its liquidity. So there is a potential demand for land that will be determined by the expectation of gains in contrast to other real and financial assets. For instance, in 1995, during the *Plano Real* when inflation was defeated, land prices fell about 40% in real terms (Plata 2001; Reydon et al, 2006)
- The demand for land in inflationary contexts is strongly related to the effect of inflation on real interest rates. If real interest rates are negative, financial assets are not attractive and, therefore, the investors will look elsewhere for real assets, such as real estate, houses, urban areas, agricultural land etc. (Reydon and Plata, 2006).
- Rural land taxes can affect price insofar as they raise the cost of maintenance. A land tax has the virtue of encouraging an increase in the productivity of idle land or where there is a low level of utilization (Reydon and Plata, 2006).
- The level of development of a country's financial system affects the price of rural land. The absence of liquidity in an economy is important because it increases the opportunity cost of money. In the case of agricultural business, with long-term investments, liquidity constraints are frequent. For example, in a country with an underdeveloped financial system, only those agents that have portfolios with highly liquid assets⁹ can purchase land. As a consequence, there will be little demand for the purchase of land, but the demand to rent land will be higher.
- Transaction costs in the land markets are the combination of several costs: bureaucracy, research, asset evaluation, management costs etc. High transaction costs in the land market are the major factor behind the low incentive to trade in land.
- Finally, the socioeconomic and political environment where trading in land takes place, is crucial. If other investments and investment opportunities are not as attractive and safe, land prices will increase in consequence of the high returns and security offered by this asset. If the legal system is complex or unstable, if there is no security in renting land and if there is an unstable political environment, no long-term investment will be done, which will affect land prices. If the ownership of the property is at risk, with invasion or expropriation for example, land prices will be negatively affected (Reydon, 1992). The entire economic, social, and political contexts of the specific land market should be taken into account when analyzing.

3 Methodology for rural land price determination in specific markets¹⁰

This item presents a methodology for determining rural land prices in specific markets in Brazil, defined as Homogeneous Zones. The Homogeneous Zones are defined by cluster analysis, using Ward's method (Ward, 1963) and SPSS software, based on the similarity of municipalities with regard to a set of characteristics: the land's agronomic condition, location, the main stakeholders in the market, level of mobility, expected purchase prices and level of urban development.

Land prices in specific markets are determined by local variables, so markets have to be analyzed using disaggregated information. To use the state or province level in Brazil would be too aggregated, so these will be divided into Homogeneous Zones, aggregating municipalities using cluster techniques. The variables used to aggregate the municipalities in order to form the Homogeneous Zone are primarily economic, social and agronomic. After the aggregation of municipalities into Homogeneous Zones, a questionnaire will be applied for each state to a random sample of recently traded properties, to capture values for the main variables which will be taken into account in the forecast land price model.

The methodology used to study rural land prices in specific or local markets observes the following stages: i) formation of a secondary database to establish Homogeneous Zones through cluster techniques, using secondary information, ii) formation of a primary database with the application of a questionnaire to the purchasers of rural properties, by stratified samples, to find real land prices and the explanatory variables, iii) statistical analysis of the primary information database to exclude incomplete or incorrect data, such as extreme values, and to obtain responses focusing on the market price equation,

and iv) create a computer program (offline) and build a database (web) to estimate land prices from information obtained from people accessing the system.

3.1 Primary information from a Homogeneous Zone (fieldwork)

The primary information for the study of land price dynamics in specific markets will be obtained through fieldwork conducted using random sampling of properties traded in a Homogeneous Zone. The sample must be distributed proportionally to the number of municipalities that make up the Homogeneous Zone. The sample has to achieve a minimum of 50 deals per Homogeneous Zone.

The cadaster of trades by municipality, used to define the random sample, consists of a list of completed deals for the respective areas, obtained from the public notary. During interviews, the researchers use printed application forms that are filled out and they get electronic codes. Another program receives the database which is analyzed and the final processing is performed. These stages are as follows: more advanced critical routines with registers being checked for duplication, extreme values, as well as several other logical processes like: price deflation, composition of data and interaction with the external database. The outcome at this stage is a database which will be used for the statistical analysis.

Trained interviewers performed the fieldwork and applied the questionnaires. The first stage of the research was carried out at the notary's office, identifying all the deals noted on the area statements of the predefined properties in the Homogeneous Zone. The purchasers, once identified, were interviewed using a 100-item

questionnaire that generated more than 250 variables. The variables cover the following types of property characteristics: physical (soil, climate, topography), productive (system of production, location, access), infrastructure of the property (fences, buildings) and expectations (regional situation, local investments). This information was input to the database to be used in the statistical analysis that defined equations for the land price determination.

3.2 Model to determine the land price in a Homogenous Zone

From the refined database, using as a minimum unit of analysis the deals completed in a Homogeneous Zone, consisting of a group of municipalities, multiple regressions were estimated to establish equations to determine land prices to be used as a basis for the forecasting of the price for a specific property. The model uses, as a dependent variable, the rural land price at a specific moment in time and, as independent variables, the farms' relevant characteristics that explain the land price in the same specific market. The estimation method is that of the ordinary least squares (OLS), with the use of the forward stepwise technique. This technique consists of the inclusion of the variables of highest explanatory power in the regression equation, which in statistical and theoretical terms, contribute to a higher level of explanation of the variation of the dependent variable. The stepwise technique permits a more parsimonious model to be attained, which will be used to predict prices while observing, however, the theoretical relationship between dependent and independent variables.

Land prices are determined by two types of variables: productive, those related to land as a production factor and speculative, those related to land as an asset that maintains value. To study the variable

effects on land prices in specific markets, from the information collected in the fieldwork, the following equations will be estimated:

$$PRICE_t = a_0 + a_1 X_{1t} + a_2 X_{2t} + \dots + a_k X_{kt} \quad I = 1, 2, \dots, k \quad t = 1, 2, \dots, n \quad (2)$$

PRICE: Price per hectare of property negotiated.

This variable can be represented by the current price (PCTE) or by the real market price (PREAL).

The latter was obtained using the current price deflated by the IGP-DI general price inflation index, base January 2004.

X_i : represents the relevant variables that explain the variation in rural land prices in the specific market. These variables can change from one Homogeneous Zone to another.

t : represents the different Homogeneous Zones.

The basic hypothesis which it is aimed to test in the model (2) is the existence of a significant relationship between the specific market and the proxy variables that capture the expectations of the buyers at the time of deciding the land price.

3.3 Updating of the model of land price determination

Whenever new property trades have been analyzed, they can be included in the sample. It will permit an improvement of the equation due to there being a larger sample. In the case of the Land Credit Program, the loan obtained to purchase the land will permit a quick improvement in the sample, making it easier to update the model. This updating will be achieved by inserting the same variable in all the new deals from the government programs: *Consolidação da Agricultura*

Familiar (Consolidation of Family Farming – CAF), *Crédito Fundiário de Combate a Pobreza Rural* (Land Credit and Poverty Alleviation Program – CF-CPR) and *Nossa Primeira Terra* (Our First Land – NPT). The addition of data will be performed through the integration of the collected program and analysis of data used in the routine PNCF (National Land Credit Program) process linked to a database on the web, which receives and stores the new information. This model update process requires a model maintenance team which will monitor data input and make the necessary adjustments to the equations, so that they reflect market changes and incorporate the new data.

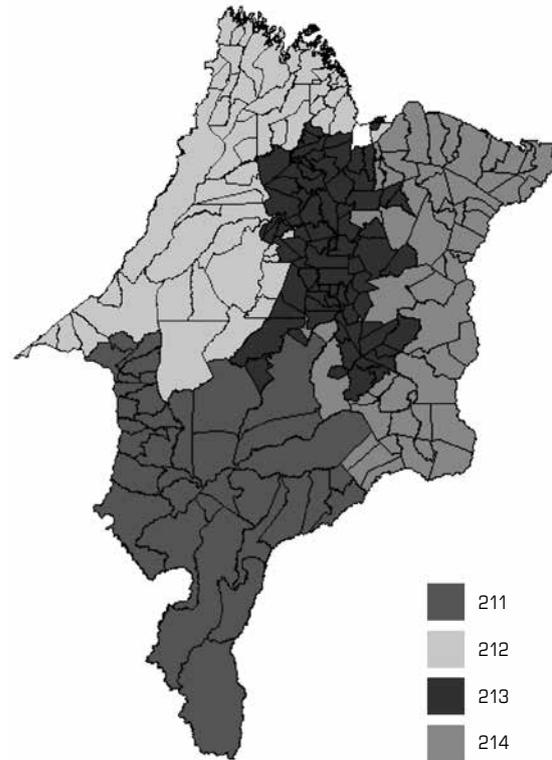
4 Application to the state of Maranhão, Brazil

This item presents the application of the hedonic land price model to the case of a Homogeneous Zone in the state of Maranhão, in northeastern Brazil. Using cluster analysis, it was possible to identify four major Homogeneous Zones, as illustrated in Figure 1. From the four Zones, the Homogeneous Zone chosen was number 211 (in red in Figure 1), with 35 municipalities. From these 35 municipalities, for the fieldwork, 75 questionnaires were collected in 8 sampled municipalities.

4.1 Refinement of the sample

In Homogeneous Zone 211, despite the strict control over the data collection process, the possibility of incorrect values had to be carefully considered. Very low or very high prices could be an indication of some kind of problem with the data. Thus, the refinement of the sample was based on the upper price limit with a 95% confidence interval. Due to the high dispersion of prices, the lower limit of the confidence interval was

Figure 1_Geographic distribution in Homogeneous Zone in the state of Maranhão, Brazil.



Source: Author, based on field survey data.

negative. Transactions with prices under R\$ 30.00 per hectare were eliminated based on a qualitative analysis of the market which indicated that such values would be extremely atypical. One observation showed a price under R\$ 30.00 per hectare, and the prices in six cases were higher than the upper limit of the confidence interval (R\$ 409.70 per hectare).

Synthesizing the 68 observations, according to the sample the average land price is R\$ 147.40, with a minimum value of R\$ 30.73 and a maximum of R\$ 376.22

and a standard deviation of R\$ 79.15. Seven cases were eliminated as outliers detected via Mahalanobis Distance, Cook Distance or Standardized residuals, leaving 61 observations in the final model.

4.2 Multiple regression model and model variables

The multiple regression model, used to explain and forecast the land price in Homogeneous Zone 211, starting

from a group of about 250 variables using the forward stepwise technique¹¹, selected 5 explanatory variables. The logarithm of land price per hectare (LNR\$/ha) was the dependent variable.

The variables that best explained the land price are those described in Table 1.

Table 2 presents descriptive statistics of the independent variables.

Table 1 Description of model variables.

Variable	Description	Expected sign of the estimated coefficient
Electricity	Dummy variable that indicates access to electricity. It has a value of 1 when the farm has access to electricity, otherwise 0.	Positive, as besides representing benefits from electricity itself, this variable may be a proxy of other characteristics of infrastructure, which usually come together with electricity.
Improvements	Dummy variable that indicates the existence of improvements on the farm, such as barns, for example. It has a value of 1 if there are improvements on the farm, otherwise 0.	Positive, since improvements increase production options.
Rock Fragments	Dummy variable that indicates the presence of rock fragments, which is considered to be good (1): soil with no mechanization restrictions due to rocks, or bad (0): soil with rock fragments that makes mechanization impossible.	Positive, since it is expected that the property, where rocks do not interfere with the use of mechanization, have higher prices. Those in which rock fragments make mechanization impossible have lower prices.
Soil	Composite index that considers soil's physical properties, such as depth and texture. This index varies in a range from 10 to 100.	Positive, as soil with better physical properties permits greater land productivity and rent.
Subsistence	Dummy variable; value 1 when the system of production of the property is agriculture and cattle-raising related to subsistence and trade of surplus, and 0 in the opposite situation.	The sign depends on the group of production systems in the Homogeneous Zone in question.

Source: Author's elaboration based on field survey data.

Table 2. Summary statistics of independent variables

Variable	Observations	Mean	Minimum	Maximum	Standard deviation
Electricity	61	0.61	0	1	0.49
Improvements	61	0.25	0	1	0.43
Rock Fragments	61	0.74	0	1	0.44
Soil	61	78.30	50	96	7.89
Subsistence	61	0.28	0	1	0.45

Source: Author's elaboration.

The mean, in the case of dummy variables, represents the fraction of cases in which the variable assumes the number 1 as value (e.g. 61% of the cases have access to Electricity).

4.3 Estimated coefficient

The regression model explains approximately 70% of the variance of the natural logarithm of land price per hectare, as can be seen in Table 3. Table 3 also shows the main statistics from the econometric model to predict the natural logarithm of the price of rural land per hectare in Homogeneous Zone 211.

Table 3 presents the value, standard error, statistic t and p-Value of the estimated coefficients¹².

Table 3. The estimation results.

Intercept/Variables	Value	Standard error	Statistic t	p-Value
Intercept	2.831	0.434	6.531	0.000
Electricity	0.293	0.085	3.442	0.001
Improvements	0.455	0.092	4.943	0.000
Rock fragments	0.450	0.104	4.317	0.000
Soil	0.019	0.006	3.292	0.002
Subsistence	-0.254	0.089	-2.852	0.006
R ²	0.70			
R ² adjusted	0.68			
F statistic	26.17			

Source: Author's elaboration.

According to Table 3, all the variables were significant to an error level lower than 1%. All the coefficients present the correct sign, as defined in Table 2.

The regression intercept indicates that, when the value of all the dependent variables is zero, the price forecast by the model is R\$ 16.96 per ha (antilog of

2.831). Because the dependent variable is the natural logarithm of rural land area, the value of B has different interpretations, which varies according to the functional forms of the explanatory variable referred to, such as described in Table 4

Table 4. Interpretation of parameters of the variables.

Explanatory variables coefficient (functional form)	Interpretation of estimated coefficient
Continuous variable	Logarithm of the rate of variation – the forecast value is multiplied by e^B for each unit change in the explanatory variable.
Dummy variable	Logarithm of the price variation factor – the forecast value is multiplied by e^B when this variable equals 1.

Source: Author's work on Gujarati (2008).

The B coefficient of the dummy variable Rock Fragments (0.450) indicates that, when they do not interfere with the mechanization of the land, the forecast price of the property is multiplied by factor $e^{0.450}$, or that it increases by 56.8%.

The B coefficient of the dummy variable Improvements (0.455) indicates that, when there are adequate improvements, the forecast price of the property is multiplied by the factor $e^{0.46}$, or that it increases by 57.6%.

The B coefficient of the dummy variable Subsistence (-0.254) indicates that, when the property is used mainly for subsistence purposes, the predicted price is multiplied by the factor $e^{-0.254}$, or that it is reduced by 22.4%.

The B coefficient of the dummy variable Electricity (0.293) indicates that, when the property has access to electricity, the forecast price is multiplied by the factor $e^{0.293}$, or that it increases by 34%.

The B coefficient of the Soil variable (0.019) indicates that an increase of one point in the soil index raises the predicted price of the property by 1.92%.

4.4 Assumptions of the linear regression model

Linear regression model estimators by OLS are BLUE (Best Linear Unbiased Estimator) when the residuals are homoscedastic and normally distributed. Moreover, to interpret consistently the estimated parameter sign and magnitude, no serious multicollinearity problems must be present.

Multicollinearity is an econometric problem difficult to avoid when working with cross-sectional data (data at a point in time, which show a global overview) and many explanatory variables. This study of determination of land price possesses these characteristics. Several practical rules have been developed to determine which way the problem affects the estimation of the model and which variable or variables cause it. Multicollinearity makes reference to the existence of linear relations between the explanatory variables in the model. The variance inflation factor (VIF) is the most frequently used indicator in its identification. VIF values over 10 are taken as a sign of severe problems.

In Table 5, the column Tolerance indicates the converse value of the inflation factor variance and therefore tolerance values below 0.1 indicate problems of multicollinearity. The tolerance is equal to $1 - R\text{-square}$, in which R-square is the coefficient of determination of the regression in which the explanatory variables in question are taken as a dependent variable and the other explanatory variables as independent variables in the new model.

As indicated in the table above, the Tolerance value for all the explanatory variables of the models is above 0.1, which indicates the absence of serious problems of multicollinearity.

Another relevant assumption of the multiple regression analysis is the normality of the residuals. The term of error of this regression model represents the aggregate effect of several variables related to the land price, which were not included as explanatory variables. Through the Central Limit Theorem, the joint distribution of such variables is normal. A bias from normality could indicate an error of specification, which means a relevant variable not included in the model. Moreover, the hypothesis tests of the linear regression model using OLS are based on a normal distribution of residuals.

Table 5 Multicollinearity indicators.

Variable	Beta	Partial correlation	Semi partial correlation	Tolerance	R-square	Statistic t	p-Value
Electricity	0.27	0.42	0.25	0.87	0.13	3.44	0.001
Improvement	0.37	0.55	0.36	0.96	0.04	4.94	0.000
Rock fragments	0.37	0.50	0.32	0.72	0.28	4.32	0.000
Soil	0.28	0.41	0.24	0.76	0.24	3.29	0.002
Subsistence	-0.22	-0.36	-0.21	0.94	0.06	-2.85	0.006

Source: Author's elaboration.

The results of the regression indicated that the null hypothesis of residual normality in the Kolmogorov-Smirnov test could not be rejected with a significance level lower than 39%. This means residuals can be considered normal (Table 6).

Heteroscedasticity occurs when, contrary to homoscedasticity, the term of error variance is not constant, a situation in which estimates via OLS are no longer efficient. One of the tests of heteroscedasticity most frequently used is the White test in which a regression is estimated where the dependent variable consists of the residuals and the independent variables are as per the original model; their squares and their cross products are the dependent variables. Under the null hypothesis of homoscedasticity, the size of the sample (n) multiplied by R^2 of the auxiliary regression, follows a distribution χ^2 with degrees of freedom equal to the number of regressors, i.e. $n R^2 \sim \chi^2$ gl. A value of this statistic, above the critical value χ^2 to a particular significance level, indicates problems of heteroscedasticity (Gujarati, 2008).

The regression of residuals with the independent variables of the original model, its square and cross product results indicated that: (i) the p-Value associated with statistic χ^2 (0.88 – squares) does not lead to a rejection of the null hypothesis with a significance level lower than 88%, which indicates the absence of heteroscedasticity (Table 7); and (ii) the p-Value associated with the statistic χ^2 (0.91 – squares and cross products) does not lead to rejection of the null hypothesis with significance lower than 91%, which indicates the absence of heteroscedasticity problems (Table 8)

Moreover, the values predicted by the model and the real values of the property, in the order of the latter,

Table 6_Residual Kolmogorov-Smirnov normality test

Results	Mean	Standard deviation	Absolute	Positive	Negative
Normal Parameters (a, b)	0.00	0.29	-	-	-
Most Extreme Differences	-	-	0.12	0.08	-0.12
N	61.00				
Kolmogorov-Smirnov Z	0.90				
p-Value	0.39				

Source: Author's elaboration.

Table 7_Testing homoscedasticity of the residuals (squares of the residues)

Intercept/Variables	Value	Statistic t
Intercept	-0.4087	-0.6812
Improvements	-0.01496	-0.5392
Soil	0.0121	0.7714
Electricity	-0.01402	-0.5414
Rock fragments	0.000613	0.01951
Subsistence	-0.001462	-0.05431
Soil ²	-7.168e-005	-0.6972
Residual sum of squares	0.400508	
Sigma	0.091345	
χ^2 (6)	2.4009 (0.8794)	
F statistic (6.48)	0.32778 (0.9191)	

Source: Author's elaboration.

Table 8 Testing homoscedasticity of the residuals (squares of the residues and cross products)

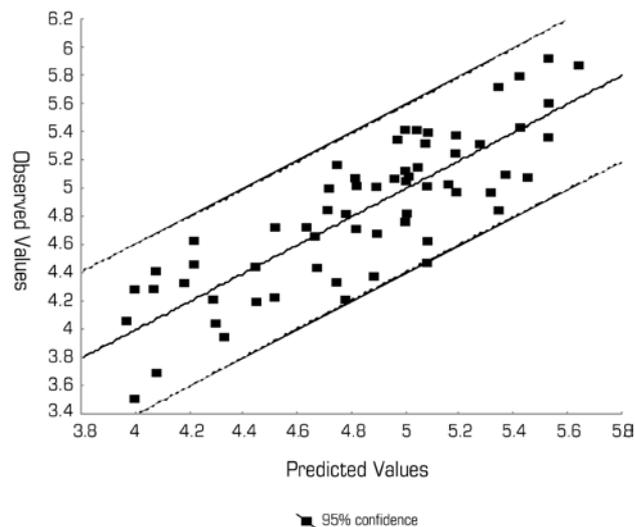
Variable/Intercept	Value	Statistic t
Intercept	-1.059	-1.049
Improvements	-0.1362	-0.2981
Soil	0.02595	0.9657
Electricity	0.4808	1.144
Rock fragments	-0.0595	-0.1008
Subsistence	-0.2996	-0.8923
Soil ²	0.0001379	-0.7284
Soil*Improvements	0.001312	0.2217
Electricity*Improvements	-0.02316	-0.3231
Electricity*Soil	-0.007114	-1.287
Rock fragments*Improvements	0.04147	0.4127
Rock fragments*Soil	-3.12e-006	-0.0004074
Rock fragments*Electricity	0.1081	1.363
Subsistence*Improvements	-0.03464	-0.3967
Subsistence*Soil	0.00384	0.8848
Subsistence*Electricity	-0.05201	-0.066
Subsistence*Rock fragments	0.008259	0.09995
Residual sum of squares	0.354851	
Sigma	0.0966343	
χ^2 (16)	9.0811 (0.9100)	
F statistic (16,38)	0.41541 (0.9692)	

Source: Author's elaboration

demonstrate the adequate adjustment of the model with all forecasts lying on a 95% confidence interval (Figure 2).

Finally, it is important to stress that this model can only be used for forecasting purposes for the range of values of the dependent variable of the database from which it was estimated.

Figure 2 Forecast versus observed price per hectare of rural land in Homogeneous Zone 211 in the state of Maranhão, Brazil.



Source: Author's elaboration.

5 Final considerations

This paper discussed and applied a methodology to explain and forecast rural land prices per hectare in specific markets. This methodology is based on a multiple regression model, with the logarithm of the rural land price per hectare as a dependent variable and, as explanatory variables, a group of variables related to physical aspects (soil, climate, landscape), production (systems of production, location, approach), infrastructure of the property and expectations (regional situation, local investments). The stepwise technique was used to select the variables included in the model: existence of Rock Fragments, Improvements, Subsistence as the main use of the property, access to Electricity and Soil characteristics.

This model provided an explanation for 70% of the variance in price per hectare of rural land. The assumptions of the multiple regression models were corroborated, in terms of the normality and homoscedasticity of the residuals. Multicollinearity was kept to acceptable levels. In general terms, the statistical, economic and econometric evaluation of the models proved to be satisfactory for the forecasting of rural land prices per hectare in the Homogeneous Zone in question. The model has been used by the Brazilian Ministry of Agrarian Development to establish limits for buying land through the different land credit programs all around the country.

Notes

¹ Homogeneous areas are groupings of municipalities based on uniform territorial characteristics. They are defined based on the homogeneity of their agronomic, economic and social characteristics, considering the productive structure, the availability of natural resources and the physical aspects of each location. It is similar to the definition given by Perroux (1967) of economic space.

² There are places, in some less developed Brazilian regions, where only subsistence is achieved and not maximum monetary gains. Primarily in these regions, extra economic factors are the dynamic feature of the land markets, for instance: tradition, line of descent, social status amongst others. Certainly, these regions, when developed

for production for the markets, demand for industrial products and with growing employment and income, will also be aimed at maximum monetary gain.

³ Any good acquired with the purpose of producing profits or that generates expectancy of a change in value, is considered an asset. This is why all goods can be treated as assets.

⁴ The level of guarantee and/or acceptance of the legal rules for the establishment of private property (legally enforced and political) are the determinants of liquidity and the dynamics of its secondary markets. In Brazil, the Land Law of 1850 established these rules generally, but because it did not have a Cadastre System and it has always been possible to regulate possession, this market

still needs significant intervention from the State.

⁵ As everything is based on expectations, there is no need for a change in the rules, if the agents think that the changes might happen; the asset price will undergo change.

⁶ The assumption of constant supply is used because it is impossible to support theoretically the existence of a supply function for an atypical asset such as land. Land cannot be produced, making it difficult to use the production theory to establish empirical supply functions.

⁷ The land's marginal productivity can also be interpreted as an opportunity cost, *ceteris paribus*, the functions of the product market and production function. This should be the value paid for the expropriation of land for land reform.

⁸ Even in inflationary environments where full indexation is present, this does not mean all prices will be equal. Therefore, it is expected that some prices will increase more than others.

⁹ These agents bought land taking into account the prices of other real and financial assets.

¹⁰ There is a need to estimate land prices in Brazil because at the notary offices the owners declare a lower value for their properties in order to pay less property transfer tax.

¹¹ A common problem in regression analysis is that of variable selection. Often, you have a large number of potential independent variables and wish to select from amongst them, perhaps to create a 'best' model. One common method of dealing with this problem is some form of automated procedure, such as forward, backward, or stepwise selection. The stepwise method is a modification of the forward-selection technique and differs from it in that variables already in the model do not necessarily stay there. As in the forward-selection method, variables are added one by one to the model, and the *F* statistic for a variable to be added must be significant at the level. After a variable is added, however, the stepwise method looks at all the variables already included in the model and deletes any variable that does not produce an *F* statistic significant at the level. Only after this check is made and the necessary deletions are completed can another variable be added to the model. The stepwise process ends when none of the variables outside the model has an *F* statistic significant at the level and every variable in the model is significant at the level, or when the variable to be added to the model is the one just deleted from it.

¹² The column "Beta" would indicate the coefficients of regression if all the independent variables had been standardized

with the average zero and standard deviation one, which permits a comparison of the influence of each independent variable on the forecast price. The following column represents the standard error of the standard coefficients. Column B indicates the estimated coefficients and the next column, its standard error.

Bibliography

- ABRAHAM, J. M.; GOETZMANN, W. N.; WACHTER, S. M. Homogenous groupings of metropolitan housing markets. *Journal of Housing Economics*, v. 3, n. 3, 186-206, 1994.
- ALSTON, L.J.; LIBECAP, G.D.; SCHNEIDER, R. The determinants and impact of property rights: land titles on the Brazilian frontier. *Journal of Law, Economics and Organization*, v. 12, n. 1, p. 25-61, 1996.
- ANGELO, C.F.; FÁVERO, L.P.L.; LUPPE, M.R. Modelos de preços hedônicos para a avaliação de imóveis comerciais no município de São Paulo. *Revista de Economia e Administração*, v. 3, n. 2, p. 97-110, 2004.
- ARRAES, R.A.; SOUSA FILHO, E. Externalidades e formação de preços no mercado imobiliário urbano brasileiro: um estudo de caso. *Economia Aplicada*, v.12, n.2, p. 289-319, 2008.
- BACHA, C.J.A. Determinação do preço de venda e de aluguel da terra na agricultura. *Estudos Econômicos*, v. 19, n. 3, p. 443-459, 1989.
- BARBIERI, A.F.; BILSBORROW, R.E. Dinâmica populacional, uso da terra e geração de renda: uma análise longitudinal para domicílios rurais na Amazônia equatoriana. *Nova Economia*, v. 19, n. 1, p. 67-94, 2009.
- BASTIAN, C.T.; MCLEOD, D.M.; GERMINO, M.J.; REINERS, W.A.; BLASKO, B.J. Environmental amenities and agricultural land values: a hedonic model using geographic information systems data. *Ecological Economics*, v. 40, n. 3, p. 337-349, 2002.
- BINSWANGER, H.P.; DEININGER, K.; FEDER, G. Agricultural land relations in the developing world. *American Journal of Agricultural Economics*, v. 75, n. 5, p. 1242-1248, 1993.
- BINSWANGER, H.P.; DEININGER, K.; FEDER, G. Power, distortions, revolt and reform in agricultural land relations. In: BEHRMAN, J.; SRINIVASAN, T.N. (EDS.). *Handbook of Development Economics*: 3B. Amsterdam, North-Holland, 1995, p. 2659-2772.
- BOWEN, W.; MIKELBANK, B. A.; PRESTEGAARD D. Theoretical and empirical considerations regarding space in hedonic housing price model applications. *Growth and Change*, v. 32, n. 4, p. 466-490, 2001.
- BRANDÃO, A.S.P. *Preço da terra no Brasil: verificação de algumas hipóteses*. Rio de Janeiro, Fundação Getúlio Vargas, 1986. (Ensaio Econômico da EPGE, 79).
- BRANDÃO, A.S.P.; REZENDE, G. The behavior of land prices and land rents in Brazil. In: MAUNDER, A. and Valdés, A. (Eds.). *Agriculture and Government in an interdependent world*. Aldershot, Dartmouth and Gower Publishing, 1989, p. 717-727.
- CASE, B.; QUIGLEY, J.M. The dynamics of real estate prices. *Review of Economics and Statistics*, v. 73, n. 1, p. 50-58, 1991.
- COTULA, L. The outlook on farmland acquisitions. Rome, International Land Coalition, 2011.
- COTULA, L.; DYER, N.; VERMEULEN, S. Fuelling exclusion? The biofuels boom and poor people's access to land. London, IIED, 2008.
- DASSO, J., SHILLING, J.; RING, A. *Real Estate*. 12th ed. Englewood Cliffs, New Jersey, Prentice Hall, 1995.
- DEATON, B.J.; VYN, R.J. The effect of strict agricultural zoning on agricultural land values: the case of Ontario's greenbelt. *American Journal of Agricultural Economics*, v. 92, n. 4, p. 941-955, 2010.
- DEININGER, K. Challenges posed by the new wave of farmland investment. *Journal of Peasant Studies*, v. 38, n. 2, p. 217-247, 2011.
- DEININGER, K.; FEDER, G. Land institutions and land markets. In: Gardner, B.L.; Rausser, G.C. (Eds.). *Handbook of Agricultural Economics*: 1A – Agricultural production. Amsterdam, North-Holland, 2001, p. 288-331.
- DELGADO, G.C. A questão agrária no Brasil, 1950-2003. In: JACCOUD, L. (Org.). *Questão social e políticas sociais no Brasil contemporâneo*. Brasília, IPEA, 2005, p. 51-90.
- EPPLÉ, D. Hedonic prices and implicit markets: estimating demand and supply functions for differentiated products. *Journal of Political Economy*, v. 95, n. 1, p. 59-80, 1987.
- FEENSTRA, R. Exact hedonic price indexes. *Review of Economics and Statistics*, v. 77, n. 4, p. 634-653, 1995.
- GOODMAN, A.C. Andrew Court invention of hedonic price analysis. *Journal of Urban Economics*, v. 44, n. 2, p. 291-298, 1998.
- GUILING, P.; BRORSEN, B.W.; DOYE, D. Effect of urban proximity on agricultural land values. *Land Economics*, v. 85, n. 2, p. 252-264, 2009.
- GUJARATI, D.N. *Basic Econometrics*. 4th ed. New York, McGraw-Hill, 2008.

- HALLAN, D.; MACHADO, F.; RAPSOMANIKIS, G. Co-integration analysis and the determinants of land prices. *Journal of Agricultural Economics*, v. 43, n. 1, p. 28-42, 1992.
- JAEGER, W.K.; PLANTINGA, A.J.; Grout, C. How has Oregon's land use planning system affected property values? *Land Use Policy*, v. 29, n. 1, p. 62-72, 2012
- KOSTOV, P. A spatial quantile regression hedonic model of agricultural land prices. *Spatial Economic Analysis*, v. 4, n. 1, p. 53-72, 2009.
- LAMBIN, E.F.; GEIST, H.J.; LEPERS, E. Dynamics of land-use and land-cover change in tropical regions. *Annual Review of Environmental and Resources*, v. 28, p. 205-241, 2003.
- LANCASTER, K.J. A new approach to consumer theory. *Journal of Political Economy*, v. 74, n. 2, p. 132-57, 1966.
- LLOYD, T; RAYNER, A.; ORME, C. Present-value models of land prices in England and Wales. *European Review of Agricultural Economics*, v. 18, n. 2, p. 141-166, 1991.
- LLOYD, T. Testing a present value model of agricultural land values. *Oxford Bulletin of Economics and Statistics*, v. 56, n. 2, p. 209-223, 1994.
- MA, S.; SWINTON, S.M. Valuation of ecosystem services from rural landscapes using agricultural land prices. *Ecological Economics*, v. 70, n. 9, p. 1649-1659, 2011.
- MSANGI, S.; EWING, M. Food, feed, or fuel? Examining linkages between biofuels and agricultural market economies. *Georgetown Journal of International Affairs*, v. 9, n. 1, p. 17-23, 2008.
- PAKES, A. A Reconsideration of hedonic price indexes with an application to PC's. *American Economic Review*, v. 93, n. 5, p. 1578-1596, 2003.
- PENDLETON, L.; MENDELSON, R. Estimating recreation preferences using hedonic travel cost and random utility models. *Environmental and Resource Economics*, v. 17, n. 1, p. 89-108, 2000.
- PERROUX, F. *A economia do século XX*. Lisboa, Herder, 1967.
- PETERS, G. Recent trends in farm real estate values in England and Wales. *The Farm Economist*, v. 11, n. 2, p. 45-60, 1966.
- PINHEIRO, F. *A Renda e o Preço da Terra: uma contribuição à análise da questão agrária brasileira*. Piracicaba, ESALQ, USP, 1980. (Tese de Livre Docência).
- PLANTINGA, A.J.; MILLER, D.J. Agricultural land values and the value of rights to future land development. *Land Economics*, v. 77, n. 1, p. 56-67, 2001.
- PLATA, L.E.A. dinâmica de preços da terra rural no Brasil: uma análise de co-integração. In: REYDON, B.P.; CORNÉLIO, F.N.M. (Org.). *Mercados de Terras no Brasil: estrutura e dinâmica*. Brasília, NEAD, 2006, p. 125-154. (NEAD Debate, 7).
- PLATA, L.E.A. *Mercados de terras no Brasil: gênese, determinação de seus preços e políticas*. Campinas, IE, UNICAMP, 2001. (Tese de Doutorado).
- REYDON, B.P. *A política de crédito rural e subordinação da agricultura ao capital, no Brasil, de 1970 a 1975*. Piracicaba, ESALQ, USP, 1984. (Dissertação de Mestrado).
- REYDON, B.P. La cuestión agraria brasileña necesita gobernanza de tierras. *Land Tenure Journal*, v. 2, n. 1, p. 127-147, 2011.
- REYDON, B.P. *Mercados de terras agrícolas e determinantes de seus preços no Brasil: um estudo de casos*. Campinas, IE, UNICAMP, 1992. (Tese de Doutorado).
- REYDON, B.P., PLATA, L.E., BUENO, A.K.; ITRIA, A. A relação inversa entre a dimensão e o preço da terra rural. In: REYDON, B.R.; CORNÉLIO, F.N.M. (Org.). *Mercados de Terras no Brasil: estrutura e dinâmica*. Brasília, NEAD, 2006, p. 207-225. (NEAD Debate, 7).
- REYDON, B.P., PLATA, L.E. O plano real e o mercado de terras no Brasil: lições para a democratização do acesso à terra. In: REYDON, B.R.; CORNÉLIO, F.N.M. (Org.). *Mercados de Terras no Brasil: estrutura e dinâmica*. Brasília, NEAD, 2006, p. 267-284. (NEAD Debate, 7).
- REYDON, B.P.; ROMEIRO, A. O mercado de terras. Brasília, IPEA, 1994. (Série Estudos de Política Agrícola, 13).
- ROSEN, S. Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy*, v. 82, n. 1, p. 34-55, 1974.
- SALLUM JUNIOR, B. O Brasil sob Cardoso: neoliberalismo e desenvolvimentismo. *Tempo Social*, v. 11, n. 2, p. 23-47, 1999.
- SANDER, H.A.; POLASKY, S. The value of views and open space: estimates from a hedonic pricing model for Ramsey County, Minnesota, USA. *Land Use Policy*, v. 26, n. 3, p. 837-845, 2009.
- TAYLOR, M.R.; BRESTER, G.W. Noncash income transfers and agricultural land values. *Applied Economic Perspectives and Policy*, v. 27, n. 4, p. 526-541, 2005.
- TRIPLETT, J. Concepts of quality in input and output price measures: a resolution of the user value-resource cost debate. In: FOSS, M.F. (Ed.). *The U.S. national income and product accounts: selected topics*. Chicago: University of Chicago Press, National Bureau of Economic Research, 1983, p. 296-311. (Studies in Income and Wealth, 47)
- WARD, J.H. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, v. 58, n. 301, p. 236-244, 1963.

E-mail de contato dos autores:

basrey@eco.unicamp.br
 ludwig@uol.com.br
 gerd@usp.br
 rafael.goldszmidt@fgv.br
 telles@iapar.br

Artigo recebido em agosto de 2011 e
 aprovado em setembro de 2012.

