LOCAL LAND MARKETS AS A SPATIAL DETERMINANT OF AGRICULTURAL SCALING

USING SPATIAL EXPLICIT FARM TYPOLOGIES

MSc. Thesis - Bart Brookhuis

Wageningen University, Chair of Land Dynamics

In collaboration with the CARE project and the Dutch "Knowledge for Climate" program

Supervised by:

dr. M.M. Bakker

dr. J.J. Stoorvogel

SUMMARY

This thesis investigates whether local land market dynamics could affect agricultural scaling. Agricultural scaling may negatively influence ecosystem services, as well as limiting possibilities for other land agents such as biodiversity conservation organizations. As farmers always buy land in their vicinity, local land markets may become locked if the spatial configuration of expanding and shrinking farms is unfavorable. Results indicate that this indeed holds true and some farmers may not succeed on the land markets in their efforts to expand or shrink.

The study area is a subjective buffer around the Baakse Beek (an IJssel-river tributary) in the “Achterhoek”. The Achterhoek is renowned for its small scale landscape. To start with, in-depth interviews were conducted. Farmers, scientists from Wageningen University, a director corporate affairs at AgriFirm (the largest Dutch farmers sillage purchasing cooperative) and the vice-chair, chairman and chairman of respectively LTO, ZLTO, LLTB (major farmers interest organizations) were asked to share their opinions concerning the regions as well as the national agricultural context, background and recent developments. Maybe even more importantly, this was done to get acquainted with actual decision making. After the theoretical background of local land markets had been set up, available census data of the Farm Structure Survey was analyzed with a multinomial logistical model using R-software. This was done to predict which farms will expand and which will shrink. After this, the willingness to pay for land as a function of distance to the buyer was constructed.

The method designed to study local land markets is also very applicable to land use and cover change modeling. It is capable of simulating declining farm numbers, but also with expansion and shrinkage of farms. The method does not rely on biased sampling (virtually all farms can be analyzed), nor does it depend on personal interpretation by the researcher or on intensions by farmers. The structure defines and observes decision making, and is able to reproduce it. It is also applicable to other regions. It makes use of existing datasets, which are generally also available for other EU member states.

The research is financed by the Dutch Ministry of Infrastructure and Environment and by the Province of Gelderland (the study area lies within the province). The study is part of CARE (Climate Adaptation for Rural arEas), which is in turn part of the Knowledge for Climate research program.

Table of Contents

1. INTRODUCTION

1.1 SOCIETAL RELEVANCE

Expanding farms have to buy their land from shrinking or stopping farms. Farmers are for various financial and socio-psychological reasons not very mobile and seek to buy land in their vicinity (e.g., see (Voskuilen and van Elk 1990)). This immobility of farms suggests that land scarcity issues can be very local and do not solve themselves(Luijt and Voskuilen 2011). For example, a potentially expanding Dutch farmer in the East corner of the country will not buy land from a shrinking farmer in the West. If there are no farmers planning to stop in the East, the potentially expanding farmer is not able to expand. It goes largely unnoticed in scientific literature that expanders and sellers might not be able to acquire or emit land at all. This has major consequences for agricultural scaling. Agricultural scaling is strongly slowed down by the scarcity of land, most probably as a consequence of a unfavourable distribution of shrinking and stopping farms towards expanding farms, Van Bruchem and Silvis of the Dutch Agro Economic Institute state (2008). To understand the consequences for agricultural scaling, but also to explore future land scarcity problems, the spatial dynamics of land markets need to be studied.

Agricultural scaling is mainly the effect of major agro-economic processes in recent history. In the past decades, the agricultural sector of developed countries did not grow as fast as the total economy did (Gylfason, Herbertsson and Zoega 1999). Consumers spent relatively less and less on food and that is why benefits for farmers in the agricultural sector have not been growing at the same extent as their costs. Also the Common Agricultural Policy (CAP) subsidies diminished and became land-based (Silvis, Oskam and Meester 2008). Due to world marketing, farmers were forced to increase their production because their commodity prices became lower(Healy and Short 1979). These processes result in expansion and intensification. This can be illustrated by the following facts. Agricultural production had risen between 1995 and 2010 with 20%, while agricultural income decreased in that timespan with nearly 25% (CBS 2010). This resulted in a decrease in the number of farms of 35% between 1995 and 2010 (see Figure 1).

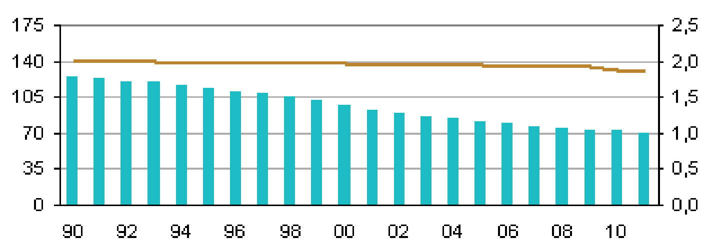


Figure 1 Number of farms in The Netherlands (left y-axis * 1000, blue bars) , and total area of agricultural land (right y-axis * 1,000,000, brown line)

To increase income, farmers may choose to expand or to intensify. If the marginal costs of expanding are higher than the marginal costs of intensifying, intensification takes place. Would it be the other way around, scaling takes place(Jager and Van Everdingen 2001). The magnitude and rate of agricultural scaling and intensification is dependent on their marginal costs, thus on prices of land. In general, scarcity of a good defines its price (Varian

2006) and land is no exception to this. At present, the magnitude and rates of these processes are barely understood, while its societal impacts can be significant. Scaling and intensification may cause ecosystem services to diminish as it effects are numerous. For example, due to landscape homogenization, agricultural scaling increases rural underdevelopment (e.g. Peat Colonies in the Netherlands), decreases biodiversity and even threatens the survival of many species (Mühlenberg and Slowik 1997). Additionally, it may undermine land purchasing practices for conserving biodiversity (Armsworth et al. 2006), decrease landscape amenity values(Lopez, Shah and Altobello 1994), may block supply of land to urbanites for rural living (Filatova, Parker and van der Veen 2009b) or may block supply of land for land reform and consolidation.

To reduce the negative effects of agricultural scaling (being environmental degradation and rural underdevelopment), the EU introduced the Rural Development Policy. This is also known as the second pillar of the CAP (Silvis et al. 2008). Its main aims "challenge the classic sectorial vision of farming as an exclusively productive enterprise" (Lowe, Buller and Ward 2002) (p .1.). Therefore farmers face new challenges. Moreover, Dutch farmers rely heavily on debts (Meulen 2008) and become increasingly older (CBS 2010). Considering these tough conditions, it can be assumed that a significant share of farms is not financially sound and a lot of them are at a 'tipping' point: if their business is in heavy weather (e.g. low prices, diseases), a large quantity does not last. As a result, comprehending farm seizure, cessation and expansion under these circumstances, can be highly relevant for appreciating agricultural scaling and intensification(van der Ploeg 2001).

1.2 SCIENTIFIC RELEVANCE

In determining the locality of land supply and demand, farmers' willingness to pay for land is a key concept. By economic laws, land scarcity will drive up willingness to pay(Varian 2006). As scarcity of land can be very local(Cotteleer et al. 2008), the willingness to pay for land is *subjective*: every farmer has unique access to land that is for sale. However, in current research, *objective* location aspects, such as distance to cities and soil quality, usually explain differences in land valuation, (e.g. (Maddison 2000, Cho and Newman 2005)). As a matter of fact, Vukina and Wossink (2000) state that, for The Netherlands, current explanations of land valuation by these objective locations aspects stand far from practice. The Netherlands have a well-developed and highly dense (agricultural) infrastructure and is probably too small for distance (to cities, nature areas) to really take effect. Technical innovations such as fertilizers reduced the importance of the biophysical environment (soil, hydrology)(Bakker, Sonneveld and Brookhuis In Prep.). Moreover, strict zoning and taxing regulations partially prevent other land uses to coerce a role on the agricultural land market. Land in The Netherlands is more or less too exploited for objective location aspects to play a role. For this reason, the significance of subjective location aspects is expected to transcend the significance of objective location aspects in land valuation in The Netherlands. This means that a farmer's perceived value of a parcel of land is mostly defined by subjective concepts as the access to parcels and the distance to the farm. Consequently, spatial land market specifications such as the spatial configuration of parcels of shrinking and expanding farms, may play a major role in farmers' decision making on farm shrinking and expansion.

Despitefully, studies on spatial land market dynamics mostly concern the peri-urban fringe (Smith, Poulos and Kim 2002, Irwin and Bockstael 2002, Filatova et al. 2009b). To some extent, rural land market dynamics have been investigated (King and Sinden 1994, Chavas and Thomas 1999). Subjective location aspects are neglected in these studies as they mostly focus on complex economic processes. Probably, the research field on subjective location aspects in The Netherland was started by Luijt (1987). He noticed a difference in farmers' willingness to pay for a parcel as a function of its agricultural land use (i.e. grassland or arable land). However, these land uses could not explain regional differences in land prices. Vukina and Wossink (2000) tried to explain these differences in prices by regions of manure surpluses (in the South and East of The Netherlands) and manure deficits. farmers buy quota to be allowed to produce manure. This can only be done by acquiring land or buying additional quota rights (from farmers in their region). Faced with scarcity, farmers were willing to pay more per hectare of land. This drove up prices of land in regions of high livestock intensity[1]. While their conceptual model predicted that farms with manure surpluses would move to regions with manure deficits, this did not happen as farmers preferably seek a solution in their vicinity. Thus, the scarcity of land depends on the density of expanding and shrinking farms.

Luijt's efforts to explain land prices recently culminated in a publication that proved the concept of local land markets (with Cotteleer and Gardebroek (2008)). They noticed that, in the Netherlands, 90 per cent of all bought parcels where acquired within 6.7 kilometres of the farm's gate. Thus, if the supply of land in a potential expander's local land market is very low, but the demand is high due to surrounding potential expanders, scarcity occurs. This can be seen from two perspectives, one is that the scarcity of land is dependent on the spatial configuration of potential expanders and shrinkers, and the other is that the success of the farms's efforts to expand or shrink relies on their spatial configuration. This is in line with the conclusions of Vukina and Wossink in 2000, who state that the scarcity of a spatial explicit good depends on the spatial configuration of the suppliers and demanders of that good . To elucidate land market dynamics it should be known who the potential buyers and sellers are. Cotteleer, Gardebroek and Luijt tried to capture the characteristics of a buyer. Unfortunately, their probit-model could only explain 5 per cent of the variance.

In summary, the land market may define developments of agricultural scaling through farm shrinkage, cessation and expansion. To study agricultural scaling this thesis analyses the spatial dynamics of land transactions by:
a) studying the spatial distribution of expanders and shrinkers, and;
b) analysing the extent to which local land markets constrain land transactions.

First, this thesis introduces a multinomial logistic model to predict potentially expanding farms *and* potentially shrinking farms from their farm characteristics. This creates the possibility to investigate the spatial location of shrinkers and expanders. To analyse the extent to which local land market could constrain land transactions, a tool that determines the willingness to pay by distance to the buyer. The study area will be a subjective area of about 300 km^2 in The Achterhoek, a typical rural area of The Netherlands.

[1] As livestock regions are based on the less productive sandy soils of The Netherlands, land on these sandy soils is even more expensive than land on more fertile clayey and loamy soils.

1.3 OBJECTIVES

The main objective of this study is:

- **To understand the consequences of local land markets for agricultural scaling by creating a model that captures and explains the spatial distribution of farms that shrink and farms that expand and their possibilities to exchange land.**

As it should be known which farms will expand or shrink in the near future in order to study local land markets, the first objective of this study is:

- To calculate the probabilities of shrinking and expanding for each farm in the study area. It is investigated how farm characteristics determine a farm's probability to expand or shrink. This is done by analyzing census data with a multinomial logistic regression model.

As the consequences of the spatial distribution of expanders and shrinkers for their possibilities to exchange land should be known, the second objective is:

- To assess whether the local land markets are still able to exchange land under the spatial distribution of expanders and shrinkers. A calculation of the willingness to pay for each parcel by each farm is performed. This is achieved by constructing a tool that quantifies willingness to pay for a parcel as a function of distance of the parcel to the expander.

1.4 THEORETICAL BACKGROUND

1.4.1 FARM TYPOLOGIES

Shrinkers and expanders could be considered farmer typologies, as shrinking and expanding is not a spontaneous decision made for the short term, but is part of a way of farming. In studies on complex socio-ecological systems (such as farms and their environment), it is very problematic to create typologies (Rounsevell, Robinson and Murray-Rust 2012). In real life, farmers show very heterogeneous behaviour even within their typologies. Contrarily, shrinkers and expanders in agricultural land markets show very homogeneous behaviour. Some theoretical background is presented here to explain the usability of these farm types and to provide a basis for furher development of a statistical model in section 2.2. Farms in the region of the Baakse Beek generally have the tendency to either cease or expand in terms of economic size units and / or hectares. In Figure 2 and Figure 3 it can be seen that the amount of farms in the Baakse Beek decline in the 'middle-class ranges'. This proves the existence (and large abundance) of shrinkers and expanders in the study area and corresponds to general conventions that scaling, intensification[2] and abandonment of farming take place.

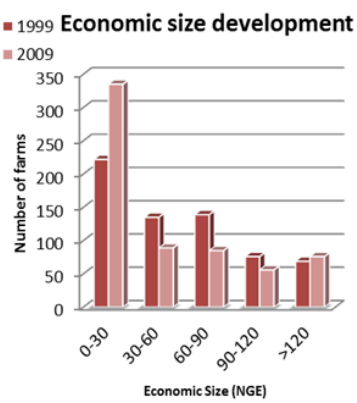


Figure 2 Number of farms per economic size class (1999 and 2009).

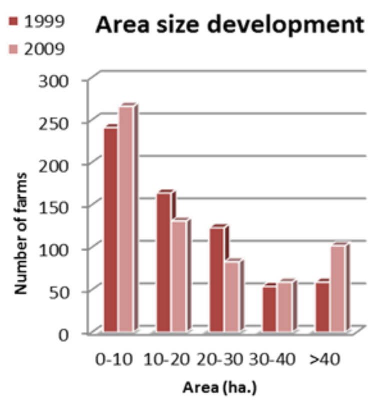


Figure 3 Number of farms per area size class (1999 and 2009)

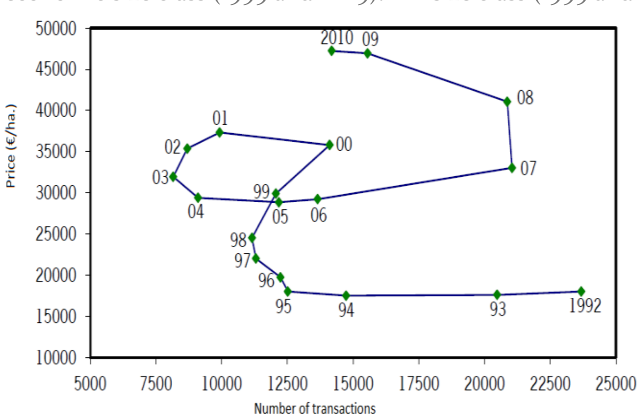


Figure 4 Market cycle of agricultural land. Price as a function of the number of transactions(Luijt and Voskuilen 2011).

In a behavioural sense, it can be said that shrinkers and expanders are both very homogeneous agent types. Within these types, farmers behave quite uniformly to market situations. A salient process that underlines this is shown in Figure 4. Figure 4 shows the market cycle of agricultural land in The Netherlands between 1992 and 2010. Between 1992 and 2001, potential shrinkers waited *en masse* for a higher price to sell their land, causing the amount of transactions to decline and the price to rise (due to scarcity) from 1995 onwards. As prices went skyrocketing, causing a bubble, most potential expanders waited for prices to drop between the year 2000 and 2004. When prices adjusted to actual market circumstances, the number of transactions rapidly increased (Luijt and Voskuilen 2011). This distinct market cycle could only have been created by very homogeneous reactions of potential expanders and shrinkers (as would the behaviour of potential shrinkers and / or expanders be hetereogeneous

---

[2]Increasing land-based production is only sustainable by farm expansion, intensification only holds for the short and mid-term Luijt, J. & M. J. Voskuilen. 2011. Grond voor schaalvergroting. 37. LEI.

within their typologies, no market cycle had occurred). This confirms that it is appropriate to discriminate farmers to expanders and shrinkers.

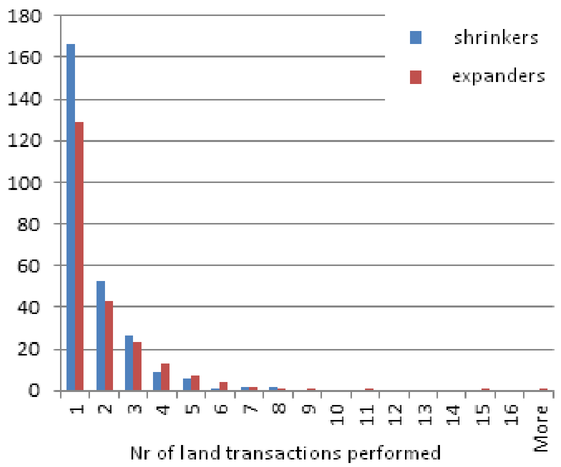


Figure 5 Frequency of nr of land transactions per farm typology

Note that these processes are not only relevant to indicate the homogeneous behaviour of potential expander and shrinker groups, but it is also relevant for assessing the fractions of potential expanders and shrinkers to all farmers. According to the Cadastral Land Sales Database, the majority of land transactions in the Baakse Beek area is performed by shrinkers. Shrinkers often just sell one parcel of land, expanders have the tendency to buy multiple parcels in a given timespan, see also Figure 5. However, the transactions represent *successful* interactions between potential expanders and shrinkers. Figure 4 can help to indicate which expander-shrinker ratio should be operated when including unsuccessful interactions of farms too. The supply and demand of land in the 1999-2009 period provided a firm basis for assessing the fraction of potential expanders and shrinkers. In this period, a land price bubble caused market stagnation between 1999 and 2005, indicating more demand than supply for land (as the prices were high enough to sell). From 2006 to 2008, most transactions occur (at a steady price), indicating a a supply and demand equilibrium. In 2009, the market headed for scarcity again (i.e., supply lags behind demand). Convincingly, this suggest that, overall, the land market had a lack of supply of land. Consequently, it is assumed that the Baakse Beek area has less potential shrinkers than expanders.

1.4.2 WILLINGNESS TO PAY

Cotteleer (2008) noticed that a parcel of land is always sold in the vicinity of the buyer. But how does this work out in practice? When is a parcel too far away from potential buyers to be sold at all? As can be seen in Figure 6, it is simply impossible to measure what the exact effect of distance to the price is, because of a lack of data points at larger distances. Whatever the distance of a parcel to a buyer, the effect of distance will not be clearly expressed in the transaction price. In practice, the *price* paid for a parcel does not express the *willingness to pay* (WTP) for that parcel for all potential expanders. This is because the willingness to pay is a subjective location aspect, (at least partly) defined by the distance to the parcel. A potential expander that is further away has a competitive deficit and will (also depending on other attributes than distance) have a lower WTP. An owner of land values his or her land also according to the subjective distance to the parcel, but this is called the *willingness to accept* (WTA). The WTA refers to which bid should be minimally placed to persuade the farmer to sell the parcel.

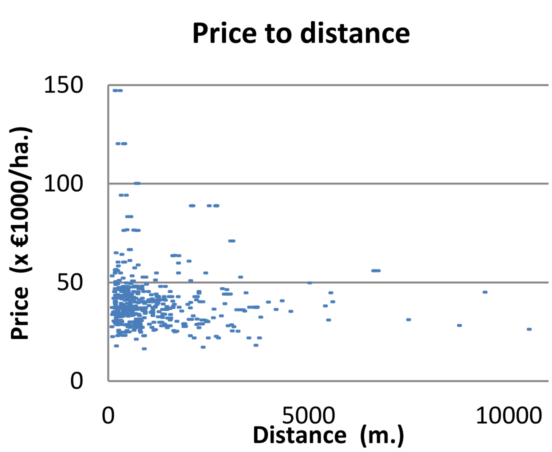


Figure 6 Transaction prices as a function of distance

1.5 STUDY AREA

The Baakse Beek area is a characteristic Dutch rural region with a relatively low population pressure and about 1000 farms. It is a subjective buffer of 285 km^2 of the IJssel-tributary the “Baakse Beek”. The region is part of the The Achterhoek which is renowned for its scenic landscape, a well-developed agricultural scene, and what Germans refer to as, a typical ‘heimat’ culture (i.e. the inhabitants tend to be proud of their region). The area (see Figure 7) is selected by the CARE project as a ‘dry’ rural area to study the area’s future resilience to climate change. However, due to its complex hydrological situation risk of flooding (in fact, due to heavy rainfall in the end of august 2010 flooding occurred), investigations for a new agricultural master plan have been started (Vorage 2010). Additionally, local governments try to buy land strategically for the purpose of water storage, nature and land reformation.

Land use mostly comprises pasture, as dairy farming is a prominent farm type. Some sugar beet (probably in rotation with potatoes and cereals) cultivation can be found. Also some fruit and three nursery farms can be found. In 1999 farms had in the Baakse Beek area had an average area of about 20 hectares, which was significantly under the Dutch average of about 30 hectares. Furthermore, economic size of the farms was also about a quarter smaller (68 NGE in the Baakse Beek area versus 89 in The Netherlands as a whole)(LEI 1999). The soil is predominantly sandy east of Hengelo and mostly clayey west of Hengelo.

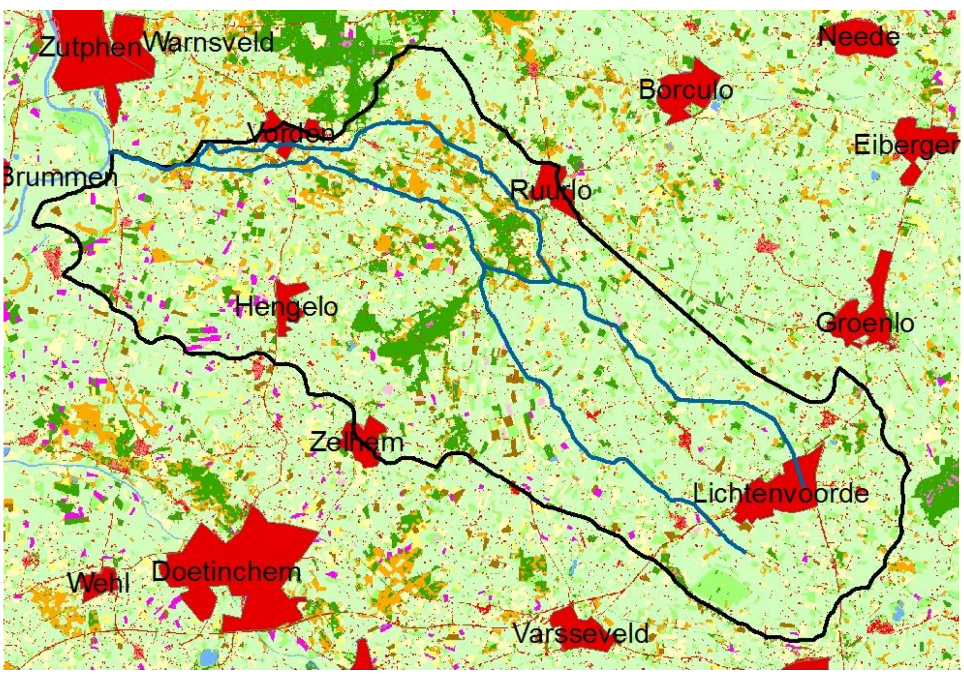
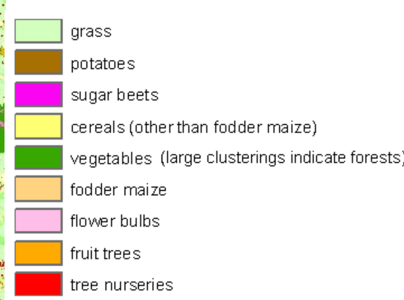


Figure 7 The black outline denotes the study area. The blue line indicates the tributary of the Ijssel known as the Baakse Beek.



The mobility of ground for agricultural purposes in the Baakse Beek area is around 60 transactions a year, according to the Cadastral Land Sales Database. Furthermore, a strength and weakness analysis by farmers at a workshop organized by LTO Noord (Agricultural Organisation of North Netherlands) pointed out the ground market is under high pressure Vorage(2010). Additionally, some more problems concerning the agricultural land market arose from the workshop:

- Parcels neighbouring the farm are too small
- High costs of parcels' management
- Bad geometry of parcels
- Too many parcels on large distance
- Small scale landscape impairs agricultural practices
- Fragmentation of agricultural area
- Little supply of land
- High prices of land
- High competition for land
- Stoppers do not sell their land
- Land is bought for other purposes than agriculture
- Penetration of the agricultural area by urbanites

In summary, farmers can be heavily affected by the local land market and the issue of land scarcity is very relevant to local farmers too.

2. MATERIALS & METHODS

By assessing local land market consequences for agricultural scaling, the spatial configuration of expanders and shrinkers should be determined. Furthermore, it should be known at what distance to the farm farmers are still willing to pay for a parcel. In the flowchart of Figure 8, the steps that have been taken to study land exchange are shown. In order to predict which farms most probably expand, shrink or remain stable an analysis with a multinomial logistic regression model is performed on census data of the years 1999 and 2009. The dependent variable is made from the difference in area size for each farm ten years after 1999, which is described in detail in section 2.1.4 on "Preprocessing". Independent variables are taken from the census data of 1999.

To assess at what distance farms are still willing to exchange land, the relationship between the willingness to pay and distance of the parcel to the expander is determined (in Figure 8, this refers to the "WTP Distance Tool"). This is discusses in section 2.4.1 and 2.4.2. Whether the WTP – distance relationship better explains actual land prices, can be confirmed by hedonic price modeling on the normal land prices and on the WTP – distance relationship corrected prices (which is elaborated on in section 2.4.3). The setup of a simulation of land exchange is treated in section 2.4.4. This basically comes down to this: if the WTP is higher than the WTA, which are dependent on distance to respectively the expander and the shrinker, a parcel changes ownership.

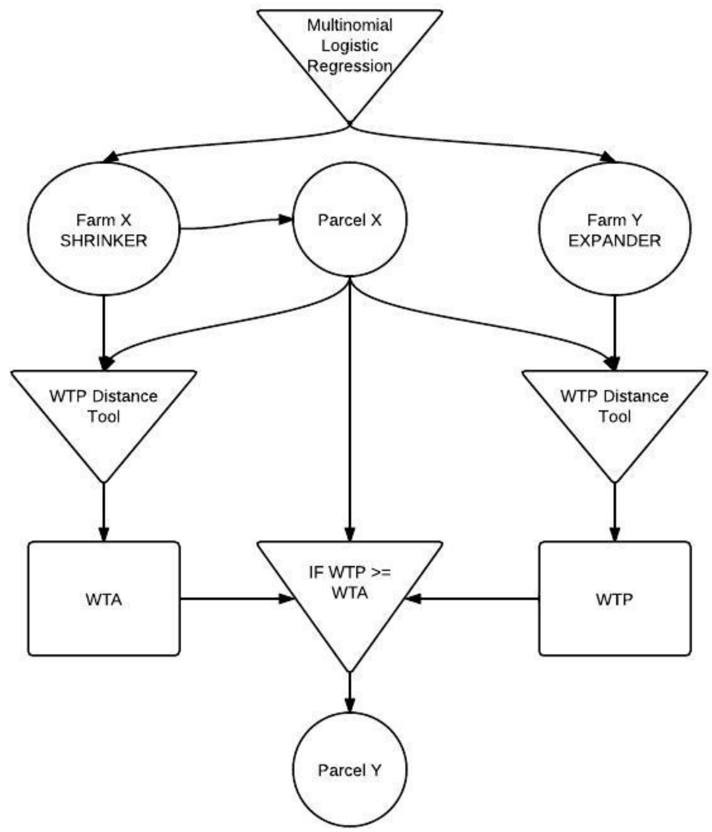


Figure 8 Flowchart of a parcel changing owner (i.e. land exchange).

2.1 DATA

2.1.1 FARM CENSUS DATABASE

The explanatory variables that control a farmer being an expander or a shrinker, represent the drivers of farm cessation, shrinking and expansion. These variables are collected from the available census data. Every year, farmers are obliged to fill out forms that give insight in their farm structure. The Dutch Institute for Agro-economics (LEI) has collected this data for about 40 years. The census contains data on age, farm size in hectares, farm type and economic size. The economic size is expressed as standard gross margin (in Dutch: Nederlandse Grootte Eenheden, NGE, and in English: Dutch Size Units) indicating the gross benefits from livestock and crops minus some specific costs (similar to the income unit used in (Cavailhès and Wavresky 2003)).

Table 1 Average, maximum and minimum values for the explanatory variables

Sample (n = 641)	*average*	*max*	*min*
age	49.6	82	24
nge	68.3	1743.5	3.33
area	19.9	574.8	0.4

The farm type is classified according to the Eurostat NEG types, indicating the land use and the grade of specialization(CBS 2009). Neglecting the vast diversity in farms, in this study all farms are split up in dairy farms and other farms. This has been done to keep the census data simple, as dairy farms make up 46% of the sample, but also there is reason to suspect dairy farming is different than most other farm types. Dairy farming is highly land-based compared to other farm types as farmers do not only need land to get rid of their manure, but also have to grow their silage(Gerbens-Leenes, Nonhebel and Ivens 2002).

Table 2 Description, units and names of explanatory variables within the study

Description	unit or value	Name
Age	Years	"age"
Area	Hectares	"area"
Economic Size Units	NGE's	"nge"
Farm Type	1 = dairy farm, 0 = other farm type	"type"

2.1.2 INFOGROMA

From the Infogroma database a farms position and unique identifier could be obtained. The unique identifier could easily be linked to the census data. This provided the opportunity to relate farms' characteristics to their geographical positions. Furthermore, the Basic Registration of Parcels(BRP), also included in the Infogroma database, is used as it also contains a unique identifier. The latest version that could be used is from the year 2008, and the earliest from the year 2001. The BRP was also used to identify the farms that bought parcels (by linking the

Cadastral Land Sales database, to the 2008 owner of the parcel). The study area has 8929 parcels, of which 8323 could be related to a farm within the study area.

The Cadastral Land Sales Database, contained in the Infogroma database, provides all transactions of importance to the Governmental Service of Land and Water management. These are transactions with a 'green' purpose(i.e. agricultural, nature and recreation). The Cadastral Land Sales Database provides the location of the transaction and the transaction price per hectare. Not all transactions could be used, as the following requirements were applied:

- the transaction must have occurred between farmers;
- no buildings were situated on the parcel;
- the transaction price was reliable (i.e. under 150.000 euro per hectare);
- the transaction took place between 1999 and 2009;
- the transactions could be linked to other cadastral data.

This resulted in 578 transactions with an average size of 3.41 hectares and an average price of €39.000 /ha.

2.1.3 OBJECTIVE LOCATION VARIABLES

For the hedonic price model, objective location aspects were used. These can be found in Table 3.

Table 3 Objective explanatory variables and their databases used for hedonic price modelling

Variable	*Map/database*
Parcel area size	BRP
Geometry of the parcel	BRP
Soil Texture null/clay/sand	PAWN Physical Properties Soil
Physical Usability for a onion, beet, grain-grain rotation	Yield loss maps
Physical Usability for grass-maize-grass rotation	Yield loss maps
Inverse distance to villages (1/(x+100))	NBLK10 Municipality database
Distance to major roads	Roadmap Major European roads
Inverse distance to ecological main structure (1/(x+100))	Provincial zoning maps
Situated within the ecological main structure	Provincial zoning maps
Ownership or leasing	BRP

2.1.4 PREPROCESSING

In order to determine the behaviour on the land markets of farms, the change in area within the census data of 1999 and 2009 has been calculated for each farm. 641 farms in the Baakse Beek area are in both datasets. In section 2.2.2 Multinomial Logistic Regression", the average parcel transaction size (3.41 hectares) is suggested to be a practical margin for area change indicating whether a farm remained stable, or expanded. Exactly 203 farms have expanded more than 3.41 hectares between 1999 and 2009. To introduce the shrinkers' market, it is assumed that there should be around 150 shrinkers: about 25 per cent less potential shrinkers than expanders. A shrinker therefore sold at least 2.3 hectares of land between 1999 and 2009, resulting in exactly 152 shrinkers.

Table 4 Absolute and relative numbers of farms within each agent type

Typology	*N (641 total)*	*Percentage (%)*
Stable	286	44
Shrinker	152	24
Expander	203	32

Now, for 641 farms a database has been built with the dependent variable being a multinomial factor with values "stable", "expander" and "stable". Independent explanatory variables are the 1999 census data variables age, area, NGE and farm type.

2.2 CREATING A LOGISTIC MODEL

To explain agent typologies by their characteristics, a regression model has been set up. But first, the construction of the logistic model is described. This is essential in understanding further model building and interpretations. A logistic regression model is often used when the outcome variable is dichotomous or binary, Hosmer and Lemeshow say in their second edition of their book "Applied Logistics Regression"(2000). Their book provides the basis for this statistical section. This section resembles some parts literally from the book, but it an attempt to keep things short, concepts are only explained briefly.

2.2.1 LOGISTIC REGRESSION

When the logistic distribution is applied, the conditional mean of the outcome variable Y given x, is written as the probability $\pi(x) = E(Y \mid x)$. The form of this multivariate logistic regression model is expressed as:

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p}}$$

This can be transformed with the 'logit transformation' (equation 2) towards a logit model that has many of the desirable properties of a linear regression model. These properties for the logit, g(x), are that it is a linear function of its covariates (x_p), and coefficients (β_p), and may range from $-\infty$ to ∞, depending on x. The logit is obtained is as:

$$g(x) = \ln\left[\frac{\pi(x)}{1 - \pi(x)}\right] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p = \beta' X'$$

Although the logit makes modeling more easy, fitting the logistic regression model to data is different from fitting a linear regression model. Linear regression uses a general method called *ordinary least squares* minimizing the sum of squared deviations of the observed values of y from the predicted values Y and leading to the *least squares* function (where $y = E(Y \mid x) + \varepsilon$, the ε represents the error term, i.e. the deviation of the observation from the conditional mean). In linear regression, these deviations are all assumed to be distributed normally and constantly. For logistic regression, this is not the case. One still may write $y = \pi(x) + \varepsilon$. However, y can only take two values: 0 or 1. If $y = 1$, then $\varepsilon = 1 - \pi(x)$ (with probability $\pi(x)$) . And if $y = 0$, then $\varepsilon = -\pi(x)$ (with probability $\varepsilon = 1 - \pi(x)$). Therefore, ε is distributed binomially with mean zero and a variance that equals $\pi(x)(1 - \pi(x))$.

As ordinary least squares does not work with logistic regression, the *maximum likelihood* measure is used. To apply the method of maximum likelihood in logistic regression, the likelihood function must be constructed. When Y is coded 0 or 1, $\pi(x)$ gives the conditional probability that Y is equal to 1 given x. That is $P(Y = 1 \mid x) = \pi(x)$ and $P(Y = 0 \mid x) = 1 - \pi(x)$. So, for an individual observed pair (x_i, y_i), the contribution to the likelihood function will be:

$$\pi(x_i)^{y_i}[1 - \pi(x_i)]^{1 - y_i}$$

The likelihood function for all observations is a product of these terms, as the observations are assumed to be independent. This gives:

$$l(\beta) = \prod_{i=1}^{n} \pi(x_i)^{y_i}[1-\pi(x_i)]^{1-y_i}$$

Mathematically (as $l(\beta) \rightarrow 0$, most computers and software cannot deal with values very close to zero) it is easier to work with the log likelihood:

$$L(\beta) = \ln[l(\beta)] = \sum_{i=1}^{n} \{y_i \ln[\pi(x_i)] + (1-y_i)\ln[1-\pi(x_i)]\}$$

The log likelihood has to be maximized to find estimates for the coefficients. This is done by differentiating this function with respect to the coefficient vector $\beta' = (\beta_0, \beta_1 \dots \beta_p)$ and set the resulting maximum likelihood equation to zero. With iterative numerical approaches these maximum likelihood parameter coefficients are then calculated (in a software package, called R 2.14.1). Whatever the coefficients or the model fit, the sum of the observed values is equal to the sum of the expected values.

Significance of the calculated coefficients can be tested by comparing the model performance of models with and without the variable in question. Performance is assessed from the log likelihood. First, the model is compared with a so-called saturated model (i.e. a model that has as much variables as it has data points) – creating a likelihood ratio:

$$l_{fit} : l_{sat} = \left[\frac{likelihood\ fitted\ model}{likelihood\ saturated\ model}\right]$$

However, the distribution of this ratio is unknown and hypothesis testing is therefore impossible. When the minus 2 natural logarithm of the likelihood ratio is applied, a known distribution is obtained. This is called the *deviance* and can be written as (by definition, the saturated model equals 1):

$$D = -2\ln[l_{fit}]$$

The deviance can be used to calculate the explained variability of the model, (just like r^2 does for linear regression). This is done by fitting a 'null' model first, only containing an intercept variable (β_0). Then the explained variability is:

$$pseudo\ R^2 = \frac{D_{null} - D_{model}}{D_{null}}$$

When assessing the significance of an explanatory variable(i.e. age, area, economic size and farm type), the deviance of both models is compared via:

$$G = -2\ln\left[\frac{likelihood\ without\ the\ variable}{likelihood\ with\ the\ variable}\right]$$

The null hypothesis then reads that, considering the coefficient vector **β'**= (β_0, β_1... β_p), that the p + 1 coefficients for the independent variables are equal to zero. The distribution of G will be chi-square (binomially) distributed with p degrees-of-freedom. For assessing the significance of an explanatory variable p within the fitted model, the Wald-statistic (t) is used. The Wald statistic is defined as

$$t = \frac{\beta_p}{SE_{\beta_p}}$$

The variable is significant when $t < -1.96$ or $t > 1.96$.

When the model has been established, interpretation of the results in logistic regression is not as straightforward as in linear regression. When the coefficients are estimated by maximum likelihood, the results can be explained by a measure of association, the *odds ratio*. Due to its relative ease of interpretation, the odds ratio is the parameter of interest. Consider the logit $g(x) = \beta_0 + \beta_1 x_1$. Then, $g(1) - g(0) = \beta_1$. The odds of the probability that x = 1 is defined by $\pi(1)/[1-\pi(1)]$. For x = 0, the odds is defined as $\pi(0)/[1-\pi(0)]$. The odds ratio is then defined as:

$$OR = \frac{\pi(1)/[1-\pi(1)]}{\pi(0)/[1-\pi(0)]} = e^{\beta_1}$$

Thus the odds of y = 1 versus y = 0 is e^{β_1} times higher. This can be interpreted as: when x_1 is increased with c units, the odds of y = c versus y = 0 is $e^{c\beta_1}$ times as high. This means that the log odds ratio equals the difference between the computed logits, that is $\ln(OR) = c\beta_1$. The log odds ratio equals β_1 when increasing x_1 with one unit.

Furthermore, the assumption of linearity in the logit function is inherent to logistic regression. Due to limited resources, it is here assumed that the relationship between the continuous explanatory variables and the logit function is linear after a natural logistic transformation of nge and area. To make things a little more complicated, *interaction* terms may be included in the model. This means that the logit is not linear anymore. Let γ be an interaction term (also called a design or dummy variable) so that the logit equals:

$$g(\gamma, x) = \beta_0 + \beta_1\gamma + \beta_2 x + \beta_3\gamma * x$$

If the design variable γ can hold two values, being $\gamma = 0$ and $\gamma = 1$, the log odds ratio of this difference is defined as:

$$\ln[OR(\gamma = 1, \gamma = 0, x)] = \beta_1 + \beta_3 x$$

The interaction term odds ratio is thus very different from a normal independent variable odds ratio as it depends on the values of the interacting variables.

Obviously, when administering logistic regression to the data, this cannot be done for all three farm typologies at once. It is then impossible to calculate to what extent an agent is a combination of all three typologies as probabilities do not up to zero. This makes it impossible to work with a binary logistical regression method.

2.2.2 MULTINOMIAL LOGISTIC REGRESSION

If the outcome variable is nominal with more than two levels, the logistic regression model can be turned into a multinomial logistic regression[3] (MNLR) model. In this study, three levels of outcome indicate the typologies (stable, buyer and shrinker). In the binary outcome model, one logit was used. In the MNLR model with four levels of outcome, three logits are needed. First, the outcome variable is denoted Y, which is coded as j = 0, 1, 2, respectively meaning stable, expander, and shrinker. As an extension, Y = 0 is denoted as the referent outcome. The logit can be written as a product of the vectors **x'** and $\boldsymbol{\beta}_j$, resulting in:

$$g_j(\boldsymbol{x}) = \ln[\frac{P(Y=j|x)}{P(Y=0|x)}$$

$$g_j(\boldsymbol{x}) = \boldsymbol{x}' * \boldsymbol{\beta_j}$$

Then the conditional probability reads:

$$P(Y=j|x) = \pi_j(x) = \frac{e^{g_j(x)}}{\sum_{j=0}^{2} e^{g_j(x)}}$$

As a consequence of this construction, $\boldsymbol{\beta}_0$ = 0 and $g_0(\boldsymbol{x}) = 0$. To construct a likelihood function, three binary variables are created, that is Y_j. The sum of these variables is always one, as only one variable 'reacts' (and turn from 0 to 1) on Y = j. Then the likelihood function for a sample of *n* independent observations is:

$$l(\boldsymbol{\beta}) = \prod_{i=1}^{n} [\pi_0(\boldsymbol{x}_i)^{y_{0i}}\,\pi_1(\boldsymbol{x}_i)^{y_{1i}}\,\pi_2(\boldsymbol{x}_i)^{y_{2i}}\,]$$

And the log likelihood, :

$$L(\boldsymbol{\beta}) = \sum_{i=1}^{n} y_{1i}\,g_1(\boldsymbol{x_i}) + y_{2i}g_2(\boldsymbol{x_i}) - \ln\left[1 + e^{g_1(x_i)} + e^{g_2(x_i)}\right]$$

The maximum likelihood estimates of $\boldsymbol{\beta}$ can be found by taking the first partial derivative of $L(\boldsymbol{\beta})$ with respect to the 2(p+1) unknown parameters and equaling these derivatives to zero. Solving these non-linear equations can be done with numerical iterative approaches in the "mlogit" package for multinomial logistic regression(Croissant 2010), and the "nnet" package for neural networking(Venables and Ripley 2002) in R-software.

The odds ratios of an outcome Y = j versus the reference outcome Y = 0 is then defined by:

$$OR_j(a,b) = \frac{P(Y=j|x=a)/P(Y=0|x=a)}{P(Y=j|x=b)/P(Y=0|x=b)}$$

Coefficient estimates and odds ratios are similarly related as in logistic regression. Significance of covariate inclusion can be calculated with the likelihood ratio test. The deviance, D, is also used here to calculate which

[3] The multinomial logistic regression model is in literature sometimes referred to as a 'discrete choice model'. 'Multinomial' can also be replaced by 'polychotomous' or 'polytomous'.

model fits best to data. Nonetheless, the explained variability cannot be described satisfactorily by some sort of pseudo McFadden R^2 (which is returned by the "mlogit" package), considering that D only gives binary information. D consists of the sum of terms that describe models probability of yielding a correct prediction. Would one draw randomly from the probability distribution, the distribution space of the second and third largest probability could also yield correct outcomes. For example, if the outcome variable is dependent on very high threshold values (e.g. to be an expander or shrinker, one must respectively emit or acquire at least 10 hectares) the model is able to predict very well those outcomes that lie between those threshold values because that range is very large. However, the range to become a shrinker or an expander is very small and is therefore based on fewer observations. This makes it difficult to predict whether farmers will expand or shrink, although the model overall performs very well (as it predicts the stable outcomes very well). If the number of agents in each outcome category are equal, a null model (a model without explanatory variables) has maximum possible likelihood compared to models that do not have equally distributed outcome categories among the agents.

Regrettably, no solution can be found in literature. This means that the best trade-off between model score log likelihood and null-model log likelihood prediction must be found. That said, 20 per cent explanation of variance by a model compared to a null-model log likelihood of 100, is preferred above a 20 per cent explanation of variance by a model compared to a null-model log likelihood of 50. By all means, the problem of variance explanation by the multinomial logistic regression model should not discourage further model development. For convenience modelling and representativeness, it is here suggested to use the average area per land transaction. For further interpretation, an assessment of the predictive power (i.e. the fraction of outcomes predicted correctly by the model) is also included in the results.

2.2.3 MODEL BUILDING

A methodological framework for building a (multinomial or binomial) logistic regression model is provided by Hosmer and Lemeshow (2000). First, an individual (univariate) analysis of each candidate explanatory variable is conducted to start the first selection of variables for the final multinomial logit model. The estimated coefficient (and odds ratio) of each explanatory variable is checked for consistency with the theoretical knowledge on farm cessation, expansion and shrinkage. If conceptually sound, Hosmer and Lemeshow (2000) suggest to keep variables that have a test probability value under .25 (based on other work that had shown that the 0.05 levels does not suffice when identifying relevant explanatory variables). After building the univariate models, the models' loglikelihoods, pseudo R^2 and Pearson's chisquare statistic were obtained. Furthermore, explanatory variables coefficient estimates, odds ratios and Wald-score test have been calculated.

After the univariate models had been set up, multivariate models were build. This is done by step forward model selection. Step forward model selection is based upon entry of explanatory variables one by one. Probability entry and removal levels of test statistics are respectively .20 and .25, according to Lee and Koval (1997). Using stepwise forward model building, the explanatory variables were added to the logit functions. Total fit of the model can be assessed through the log-likelihood or deviance, but also through the Pearson Chi-square statistic and the

McFadden R^2. Despitefully, these measures of goodness-of-fit have their limitations as they are particularly designed for binomial logistic regression. After fitting the data to the MNLR models, the best model is used to assess individual explanatory variable importance. Variables that are not significant should be excluded from the model. However, as this is operationalized, the other estimates should be paid attention to. If the removal of a variable induced large changes in other estimate(s), it is possible that the removed variable is a confounder of the other explanatory variables and careful consideration should be taken whether this variable must be maintained in the model.

Interactions between covariates are included in the model if they are conceptually sound and provide a better model fit. Interpretation of these interactions, however, can be difficult (e.g. in the case of two continuous variables). Therefore, possible interactions are also plotted with the R "Effects" package. After the interactions have been investigated, the final model is created. The final model is assessed with measures of goodness fit such as the log likelihood, R^2 , Pearson's χ^2 and the predictive power.

2.3 INVESTIGATING LAND MARKET DYNAMICS

2.3.1 WILLINGNESS TO PAY

The price paid for a parcel does not represent all farmers' willingness to pay as he or she has a subjective way of valuing the parcel, according to his distance to the parcel. Due to a lack of data at larger distances, the relationship between distance to a potential expander and the willingness to pay cannot be observed. This section tries to explain how this is dealt with in this thesis. Nevertheless, the WTP (instead of the actual transaction price) of a parcel to a farmer may also be assessed through looking at how many parcels have been bought within a certain distance to the farmers. From the Infogroma database and the BRP database, the distances of every transaction to its buyer were calculated. This resulted in the cumulative relative frequency distribution of Figure 9 and the normal frequency distribution of Figure 10.

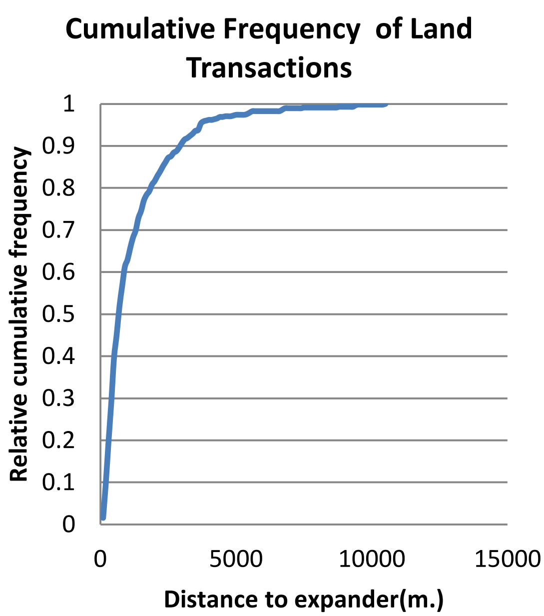


Figure 9 Cumulative frequency of land transactions as a function of distance.

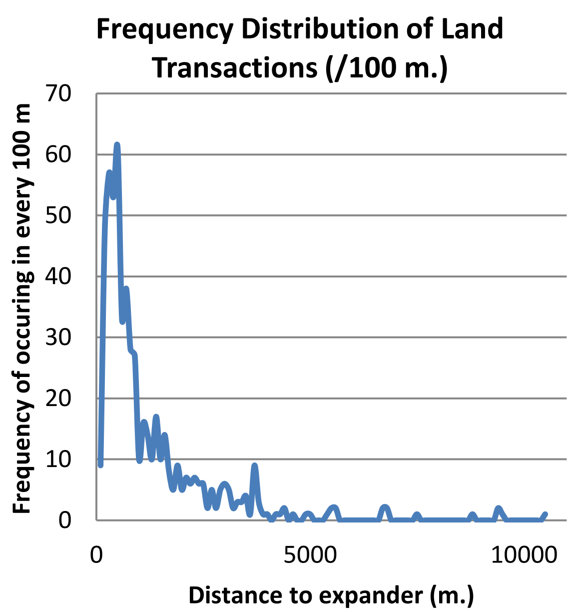


Figure 10 Frequency of land transactions as a function of distance

The frequency of occurrence for land transactions as a function of distance in Figure 10 is typically described by a Poisson distribution. But how can the frequency of land transactions be translated into a WTP-distance relationship? Logically, the single farmer decides his WTP for a parcel along a line towards the parcel of interest (one dimensional). Thus, the WTP declines linearly[4] as a function of distance(Alonso 1964). This will be explained here shortly. All linear WTP (and WTA) lines, combined in a specific spatial distribution of expanders and shrinkers, define the amount of successful transactions (which, essentially, makes up the Poisson frequency distribution). The amount of successful transactions at a given distance range is defined by the *supply of parcels* at the given distance range, and the *chance of having a successful transaction* when a certain price is offered. The offer of parcels within a given distance, r, increases quadraticly by $\pi * r^2$. The chance of having a successful transaction is linearly related to the WTP and WTA. This can also be visualized with Figure 11. In the green parts at the boundary of the circle, the chances of a successful transaction will be lower (as the WTP will be low), but the offer a parcels will be higher (which increases the chance of a parcel having a low enough WTA). The chance of having a successful transaction is linearly proportional to the WTP, and inversely proportional to the WTA. Ideally, the tool should be able to reproduce the shape of the amount of transactions in Figure 10 on page 22.

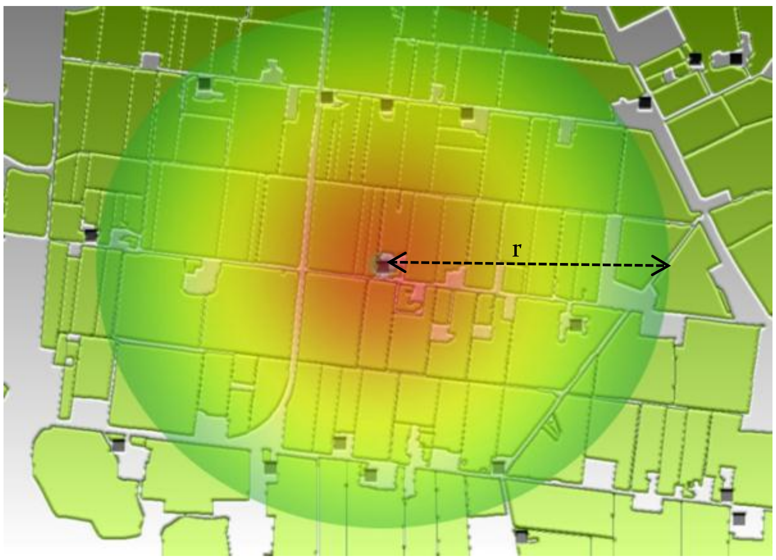


Figure 11 Visual representation of the linearly declining willingness to pay. Red indicates a high WTP, green a low. Black dots indicate farms, green fields indicate parcels.

As described earlier, the amount of successful transactions at a given distance range is defined by the offer of parcels at the given distance range, and the chance of having a successful transaction when a certain price is offered. This results in:

$$Q(\mathrm{x}) = \int p(x) * \int c(x)$$

Where:

p(x) = Offer of parcels at distance x

c(x) = Chance of successful transaction at distance x

Q(x)= Quantity of successful transactions within distance x

Then, p(x) is defined as:

[4] The linear decline as a function of distance did not fall from the sky. Opposed to assumptions of an *inversed distance* hedonic value in literature, based on location interactions with the distance described by Alonso's theory of (urban) land rent, there is no interaction here as probably there are only linear cost functions involved.

$$p(x) = a * b * 2\pi x$$

Where:

a = average parcel density per m^2 (10 per square km is assumed here)

b = supply of parcels to the land market as a fraction of the total number of parcels

Then, c(x) is defined as:

$$c(x) = a * WTP\ (x)$$
$$c(x) = a * -WTA\ (x)$$

In order to reproduce the amount of transactions for different distance ranges, p(x) and c(x) need to be *integrated over x*.

To give a small summary of the above two alineas: 95% of all parcels are sold within 3,7 km of the buyer, suggesting that distance to buyer is a very important in the bundle of services a parcel of land delivers. However, this cannot be shown in actual price data because the pressure of potential buyers closer to the parcel, having a competitive lead and therefore most probably a higher WTP, is almost everywhere too high. This means that it is impossible to do a hedonic price regression. However, the frequency of occurrences of land transactions within a given distance, suggests a *linear relation* between distance and the potential buyers' WTP. Once the WTP line has been established, hedonic price theory will define if the WTP by distance tool contributes to a better explanation of land valuation.

2.3.2 WILLINGNESS TO PAY CONSTRUCTION

The relationship between distance to a potential expander and the willingness to pay cannot be observed due to a lack of data at larger distances. However, this thesis argues the relationship is linear. Hence, it is tried to construct a relationship between WTP and distance. The linear interpretation of the willingness to pay as a function of distance is constructed by the virtual lines a, b, c and d:

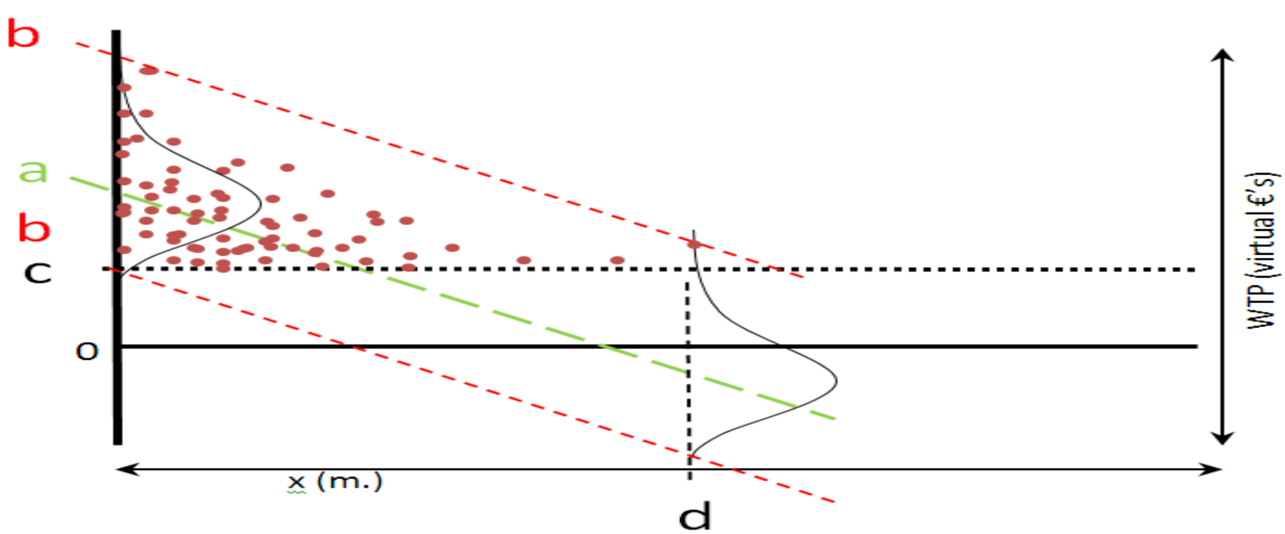


Figure 12 Hypothetical construction of the willingness to pay (per hectare).

a This is the linear function of WTP as a function of distance ($WTP^{distance}$).

b These two lines represent the lower 0.01 quantile and the higher 0.99 quantile of the Gamma distribution across distance. In the first 500 meter, it is assumed that the distribution of price (i.e. a Gamma distribution, skewed towards higher prices probably because of speculation) resembles the distribution of the $WTP^{distance}$. It is furthermore assumed that the WTP is distributed equally throughout distance. The line connects the median of the Gamma distributions and will therefore look like:

$$WTP^{distance}(x) = Gamma_{median}\,(x) = Gamma_{median}^{x=0} + \left(\frac{dGamma_{median}}{dx}\right) * x$$

c This is the lowest price for a sellers WTA for a parcel at a distance to the buyer of x = 0 metre. This is quantified by the 0.01 quantile of the estimated Gamma distribution at x = 0 meter.

d The distance wherein 99 per cent of all transactions take place. The same Gamma distribution is assumed here, with the transaction intersecting the 0.99 quantile. It is assumed that the 0.99 quantile line intersects line c at the point where 99% of all transactions occur(also defined by a gamma distribution, but more poisson like). At this distance it is assumed that the minimum price is the maximum WTP. Using the same Gamma distribution already known for the first 500 meter, the median of the Gamma distribution at this point can be calculated to construct the linear $WTP^{distance}$ function.

Hence, the (homoscedastically distributed) Gamma distribution around the WTP needs to be constructed. The transactions in the first 500 meter (n = 228), representing the distribution at 250 meter, are used to fit to a Gamma distribution. In the first 500 meter, the Gamma distribution can be defined as:

$$f(x, \alpha, \gamma) = p^{\alpha-1}e^{-\gamma p}$$

Where:

- p = price in €
- α = shape of the distribution
- γ = rate of the distribution

The shape, α, and the rate, γ, can be obtained by fitting the data to the Gamma distribution in R, using the MASS-package (Venables and Ripley 2002).

2.3.3 HEDONIC PRICE MODELING

After constructing the WTP - distance relationship other variables that may play a role within the local land market could be assessed to confirm the WTP - distance relationship is constructed correctly. This can be done by constructing a Hedonic Price Model on the actual land prices and on the error terms (the deviation of the tool to actual land prices paid) of the WTP - distance relationship. Would the WTP - distance relationship represent reality, fitting the hedonic price model on the error terms should explain more variance than fitting the hedonic price model on actual land prices. Real estate value can be estimated via Hedonic Price Theory(Rosen 1974). In other words, "*the Hedonic Price Theory is based on the concept that an implicit market for a particular quality or attribute (including geographical location) is imparted to real estate (land values and housing) by buyers and sellers* (i.e. expanders and sellers) *in a competitive market*"(Metz, Morey and Lowry 1990). In the case of agricultural land selling and buying, an expander perceives a parcel of land as a bundle of characteristics and values it according to his utility-bearing characteristics (although in this thesis this is assumed to be equal for all agents and only dependent on distance). It is an indirect method to analyse market data: it echoes how people actually behave in the market of real estate. For the past centuries, it was the main paradigm for non-survey based research towards markets of real estate. This theory has been developed from the philosophy that people behave as pleasure-seekers.

The following set of variables is used in this study:

▪ Parcel area size	(area)
▪ Geometry of the parcel	(geom)
▪ Soil Texture null/clay/sand	(as.factor(texture)0/ 1/ 2)
▪ Physical Usability for a onion, beet grain-grain rotation	(arable)
▪ Physical Usability for grass-maize-grass rotation	(dairy)
▪ Inverse distance to villages (1/(x+100))	(invdistvil)
▪ Distance to major roads	(road_dist)
▪ Inverse distance to ecological main structure (1/(x+100))	(invdistehs)
▪ Situated within the ecological main structure	(as.factor(ehs_binary)1)
▪ If parcel is leased out or in property	(as.factor(property)1)

Hereafter, the two Hedonic Price Models are:

$$P_i - WTP^{distance_i} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p = \beta' X'$$

and,

$$P_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p = \beta' X'$$

Where:

P_i	= Actual price of transaction *i*
$\mathrm{WTP}^{\mathrm{distance}}{}_i$	= Modelled price at distance *i*
β'	= Vector of model estimates
X'	= Vector of model covariates (e.g. distance to villages, zoning, density of expanders and shrinkers)

If the hedonic price model is better able to explain the error terms of the WTP - distance relationship than the actual land prices, the WTP - distance relationship contributes significantly in explaining land prices. If the error terms of the relationship do not yield better land price explanations, the relationship is not able to translate willingness to pay into actual land prices (but still represents willingness to pay by distance).

2.3.4 LAND EXCHANGE

As pointed out in the Introduction (section 1.1), transactions of land parcels can only occur if there is a supply and demand of land parcels. By modelling the spatial availability and claim of land as a function of the spatial distribution of potential shrinkers and potential expanders, the local land market situation can be simulated. The first variable of interest is the percentage of potential expanders that succeed by acquiring a piece of land. The second variable is the percentage of potential shrinkers that succeed by selling a piece of land will be calculated.

As databases of parcels in the year 2008 and the 2009 Infogroma spatial distribution of farms are the most complete databases, cencus data from the year 2009 will be used to run the model. Of all 8929 parcels in the area, 8381 could be linked to a farmer within the area. Firstly, Euclidian distances of all farms to all parcels will be calculated, using their x and y data, resulting in the farms - parcels distance matrix:

$$M_{distance} = \sqrt{(Fx^i - Px^j)^2 + (Fy^i - Py^j)^2}$$

Where:

Fx^i = Farm i's x position

Fy^i = Farm i's y position

Px^j = Parcel j's x position

Py^j = Parcel j's y position

i = 1,2,3…,959

j = 1,2,3…,8381

The matrix containing the willingness to pay as a function of distance is then created by:

$$M_{WTP^{distance}} = WTP^{distance}(M_{distance})$$

The parcels that belong to shrinkers are now set up for sale with a *willingness to accept that equals* their personal *willingness to pay* as a function of distance. Also, a situation in which shrinkers only offer one parcel (for which they perceive the lowest WTA) is studied. If an expander has a higher WTP, a transaction occurs. If there are several matches for one parcel, the highest WTP 'wins'.

Unsuccessful farms will be plotted against a probability field (achieved by spatial interpolation of the census data with Kriging in Arcgis) of their opposite farm type to reveal 'preferred' positions of unsuccessful farms in the land-market landscape.

3. RESULTS

3.1 THE MULTINOMIAL LOGISTIC MODEL

3.1.1 UNIVARIATE ANALYSIS OF EXPLANATORY VARIABLES

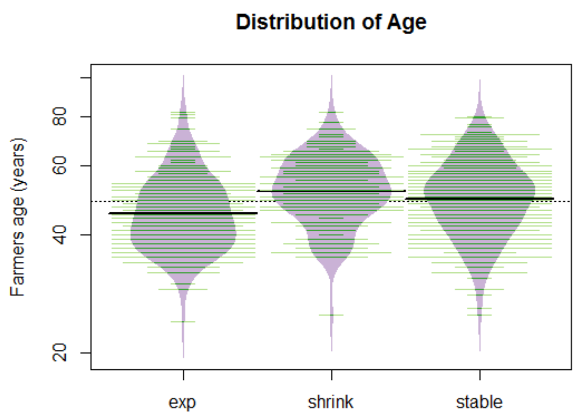


Figure 13 Distribution of age across farm typologies

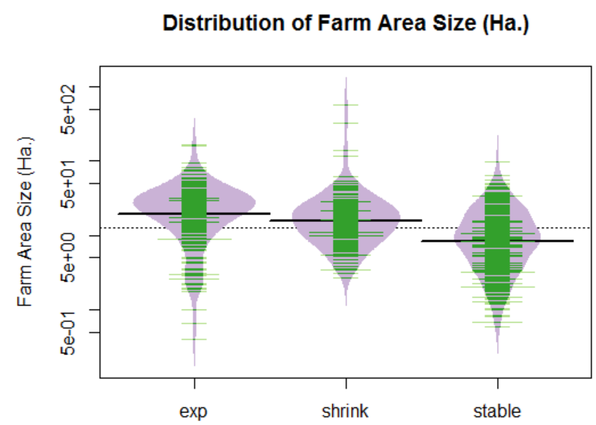


Figure 14 Distribution of farm area size across farm typologies

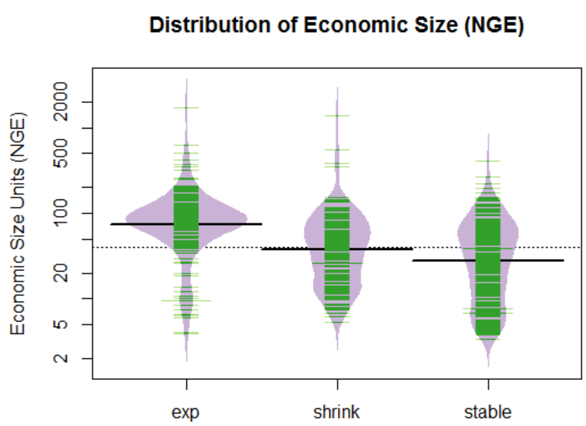


Figure 15 Distribution of NGE across farm typologies

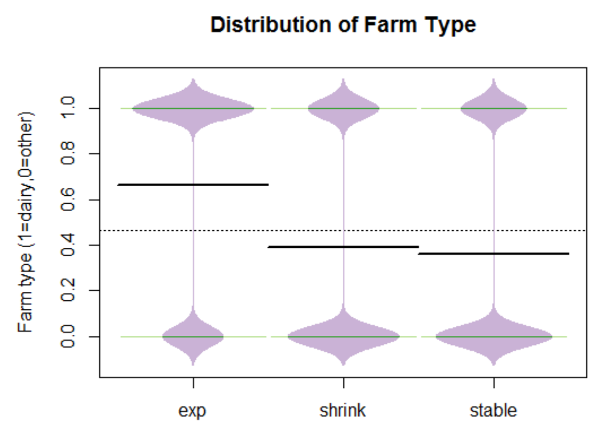


Figure 16 Distribution of dairy farms (1) across farm typologies

Figure 13 shows the age of the agent typologies. Expanders have a tendency to be younger, shrinkers tend to be a little older than stable agents. Expanders and shrinkers have on average more hectares than stable agents (Figure 14). Besides that, expanders show more activity in the lower part of the distribution. Shrinkers show more activity in the higher part of its distribution. NGE in Figure 15 shows more or less the same pattern, except for the fact that the averages are more distinct now. Expanders had clearly a higher NGE in 1999 than shrinkers, and shrinkers clearly had a higher NGE than the stable typology. In Figure 16, it is difficult no to see that most expanders consists of dairy farmers.

All univariate models are statistically significant (see Table 5). (see Table 6). They show that all variables are significant at least to one logit.

Table 5 Univariate models' measures of goodness of fit

Variable	Log-Likelihood	McFadden Pseuso R	Pearson Chi square statistic	Probability
Age	-665.46	0.03	35.05	2.46E-08
LN(area)	-632.85	0.07	100.26	2.22E-16
LN(NGE)	-625.43	0.08	115.1	2.22E-16
Type	-658.26	0.04	49.43	1.84E-11

Table 6 Univariatte coefficient estimates for each agent type logit. *For all tables holds: Significant at the 0.1, 0.05, 0.01 and 0.001 is indicated by respectively ".", "*" "**", "***".*

Variable - Logit	Estimate	Std. Error	Odds Ratio	Wald score (t)	Pr(>|t|)	Significance
Age - expander	-0.041	0.0093	0.96	-4.39	1.13E-05	***
Age - shrinker	0.018	0.0093	1.02	1.90	0.057	.
LN(area) - Expander	1.00	0.12	2.73	8.45	2.20E-16	***
LN(area) - Shrinker	0.74	0.12	2.09	6.18	6.30E-10	***
LN(NGE) - Expander	1.10	0.12	2.99	9.14	2.20E-16	***
LN(NGE) - Shrinker	0.28	0.10	1.33	2.86	0.0042	**
Type - Expander	1.26	0.19	3.53	6.53	6.67E-11	***
Type - Shrinker	0.12	0.21	1.13	0.58	0.57	

3.1.2 MULTIVARIATE ANALYSIS OF EXPLANATORY VARIABLES

Stepwise forward modelling resulted in 4 models. In Table 7, the n-th model is compared to the (n-1)th model. This shows (by means of the Likelihood-Ratio test) that each variable has clear added value to the total model. Model 4 is significantly better than the other models and will therefore be used in the next step.

Table 7 Stepwise forward model building with measures of goodness of fit for each model.

	Age	Area	NGE	Type	LogLikelihood(df)	LR-test to previous (χ^2)	Probability
Null-model					-682.98		
Model 1	x				-665.46 (4)	35.05	2.5e-08
Model 2	x	x			-616.40 (6)	98.11	<2.2e-16
Model 3	x	x	x		-589.71 (8)	53.38	2.5e-12
Model 4	x	x	x	x	-579.05 (10)	21.32	2.4e-05

As stepwise forward model selection progressed, the coefficient estimates changed. A small change in coefficient estimates is logical, major changes indicate confounder effects. The coefficient estimates are shown in Table 8.

Table 8 Coefficient estimates for each model created with stepwise forward model building.

	Coefficient Estimates				Probability			
Expanders	*Age*	*Area*	*NGE*	*Type*	*Age*	*Area*	*NGE*	*Type*
Model 1	-0.041				1.1e-05			
Model 2	-0.040	0.97			5.7e-05	2.2e-16		
Model 3	-0.037	0.47	0.81		0.00041	0.00025	7.6e-09	
Model 4	-0.036	0.31	0.88	0.59	0.00064	0.037	2.2e-09	0.015
Shrinkers	*Age*	*Area*	*NGE*	*Type*	*Age*	*Area*	*NGE*	*Type*
Model 1	0.018				0.057			
Model 2	0.019	0.76			0.053	4.5e-10		
Model 3	0.017	1.15	-0.36		0.081	1.2e-08	0.026	
Model 4	0.017	1.20	-0.29	-0.58	0.078	1.3e-09	0.058	0.017

At first sight, the shrinkers "area" coefficient estimate seems very large compared to that of the expanders. Secondly, the introduction of NGE into the model (model 3) caused a major decline in the expander area coefficient estimate of model 2. This indicates that "area" is a confounder of "nge". This implies that the area-effect diminishes because it is already partly included in the economic size unit.

3.1.3 THE PRELIMINARY MODEL

All coefficient estimates (see Table 8) in model 4 are plausible and agree with the conceptual model of shrinking and expanding farms. Aging increases the chance of shrinking, while young farmers most probably expand. A large farm size increases the chance of being an expander, but even more to be a shrinker. Economical heavyweights have an increased chance of farm expansion. An economically small farm increases the chance of shrinking. Dairy farmers have a high probability to expand, and a lower probability to shrink (bear in mind that this is always compared to stable farms).

The next three tables show the overall fit of the preliminary model. Table 10 and Table 11 have been added to show that there is some valuable model information in the second largest probabilities too. Would the most-common approach be used, 57% of all outcomes would be predicted correctly. However, 20% of the shrinkers would be predicted correctly. While the 2nd highest probability values for each agent would predict shrinkers better, overall fit would be considerably worse.

Table 9 Measures of goodness of fit for the preliminary model

	Log Likelihood	*McFadden R^2*	*Pearson (χ^2)*
Model 4	-579.05	0.15	207.86 (df=10)

Table 10 Predictive power (i.e. the percentage of observations predicted correctly)

	Overall Pred. Power	*Stable Pred. Power*	*Expanders Pred. Power*	*Shrinkers Predictive Power*
Model 4	57 %	70 %	67 %	20 %

Table 11 Predictive power in the second choice 'mode'. The second choice is the second highest probability of agent type for every farm.

	Overall 2nd choice Pred. Power	*Stable 2nd choice Pred. Power*	*Expanders 2nd choice Pred. Power*	*Shrinkers 2nd choice Predictive Power*
Model 4	28 %	27 %	19 %	40 %

3.1.4 INTERACTIONS

All possible interactions were checked and are presented in Table 12. These results are compared, by the likelihood ratio test, to the preliminary model 4. A conceptually sound and very significant interaction is the farm type – nge interaction. It is expected that economically large dairy farms have a higher expectancy to be an expander than economically large non-dairy farms as dairy farming is highly dependent on their area size (see also . The other two interactions (age-nge and nge-area) do not have a satisfying explanation.

Table 12 Possible logit interactions within the model

	age	*area*	*nge*
age	-		
ln(area)	0.095	-	
ln(nge)	0.027*	0.029*	-
type	0.57	0.088	0.0012**

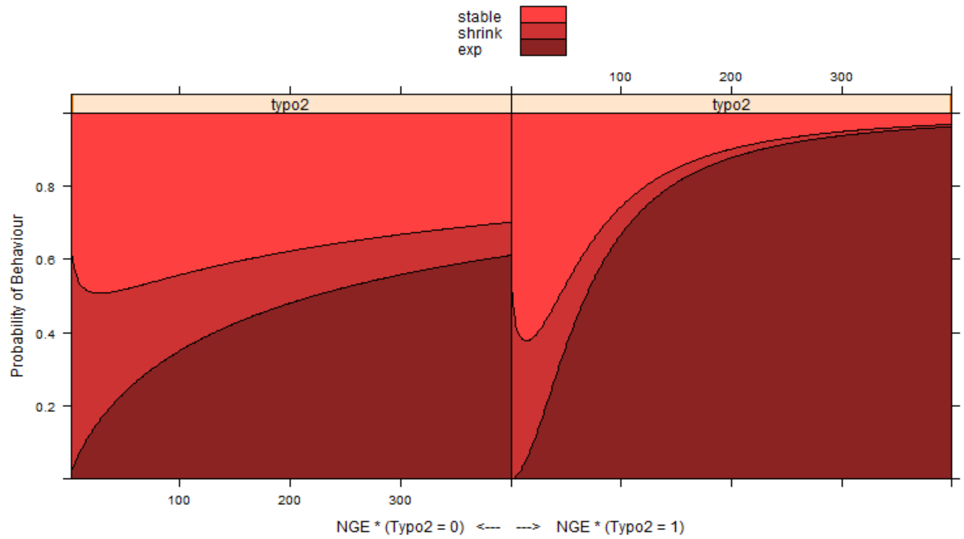


Figure 17 Visual representation of the interaction effect between NGE and farm type. Left shows the probabilities for each agent type for non-dairy farms, right shows the probabilities for dairy farmers.

3.1.5 THE FINAL MODEL

With the inclusion of the interaction between *nge* and *type*, the most accurate model is achieved. In Table 13, the coefficient estimates, their standard errors, Wald statistics, odds ratios and significance are presented. For the expanders logit, all explanatory variables except the area are significant at the $\alpha = 0.05$ level. The shrinkers' logit shows a strongly different pattern of significant explanatory variables. At the $\alpha = 0.10$ level, all variables are significant *except farm type*. This means that the farm type is not able to exclude shrinking farms from stable farms. The odds ratios are in line with the general knowledge on farm shrinkage and expansion.

Table 13 Explanatory variables and their coefficient estimates, standard errors, Wald statistics and probabilities.

Explanatory Variable	*Coefficient Estimate*	*Standard Error*	*Wald statistic (t)*	*Odds ratio*	*Probability (>t)*	
Expander						
intercept	**-2.23**	0.74	-3.01		0.00	**
Age	**-0.03**	0.01	-3.15	0.097	0.00	**
ln(area)	**0.20**	0.15	1.38	1.23	0.17	
ln(nge)	**0.68**	0.15	4.62	$1.97 + e^{(1.09 * \text{type})}$	0.00	***
ln(nge)*type	**1.09**	0.34	3.22		0.00	**
type	**-3.81**	1.38	-2.76	$0.02 + e^{(1.09 * \ln(\text{nge}))}$	0.01	**
Shrinker						
intercept	**-3.23**	0.71	-4.57		0.00	***

age	**0.02**	0.01	1.82	1.02	0.07	.
ln(area)	**1.20**	0.20	5.88	3.30	0.00	***
ln(nge)	**-0.32**	0.16	-2.01	$0.72 + e^{(0.02*type)}$	0.04	*
ln(nge)*type	**0.02**	0.27	0.07		0.94	
type	**-0.55**	1.05	-0.52	$0.57 + e^{(0.02*ln(nge))}$	0.60	

In Table 14 it can be seen that the Log Likelihood and the McFadden R^2 have slightly increased compared to model 4 at section 3.1.3. In Table 15, it can be seen how this works out. The model does not predict shrinkers or expanders better, but is able to predict the largest group of agents better by 4%. Again, like in section 3.1.3, the second choice has far more predictive power for the shrinkers.

Table 14 Measures of goodness of fit of the final model

	Log Likelihood	*McFadden R^2*	*Pearson (χ^2)*
Final model	-572.35	0.16	221.26 (df=12)

Table 15 Predictive power of the model.

	Overall Pred. Power	*Stable Pred. Power*	*Expanders Pred. Power*	*Shrinkers Predictive Power*
Final model	58 %	74 %	62%	20 %

Table 16 Second choice mode predictive power.

	Overall 2nd choice Pred. Power	*Stable 2nd choice Pred. Power*	*Expanders 2nd choice Pred. Power*	*Shrinkers 2nd choice Predictive Power*
Final model	27 %	22 %	22%	45%

With the multinomial logistic regression model constructed, a map of potential shrinkers and expanders has been created. Some areas with low expander or shrinker densities can be discerned. The amount of farms within each agent type can be found in Table 17

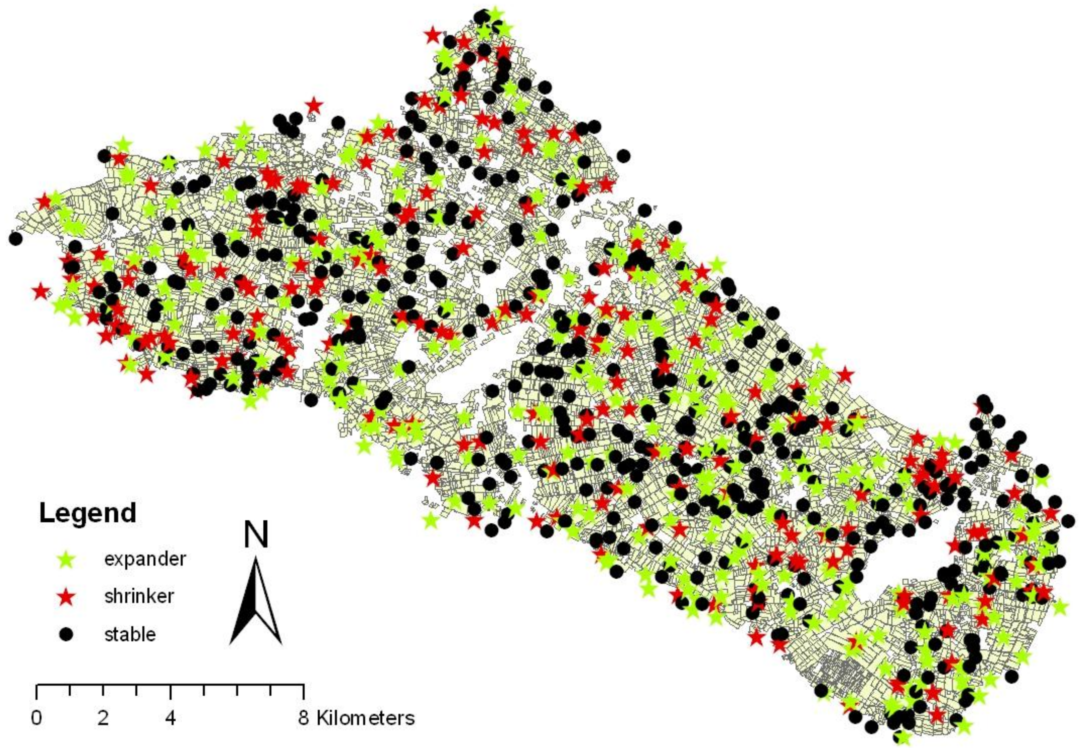


Figure 18 A map of potential shrinkers, expanders and stable farms for the year 2009.

Table 17 The number of farms within for each agent type in the year 2009.

Typology	*N (959 total)*	*Percentage (%)*
Stable	468	49
Shrinker	204	21
Expander	287	30

3.2 LAND MARKET DYNAMICS

3.2.1 WILLINGNESS TO PAY - DISTANCE RELATIONSHIP

As described in section 2.3.2, the willingness to pay relationship is constructed. The line that represents the minimum price ever to be paid for a parcel, whatever the distance, intersects the line of maximum $WTP^{distance}$ (i.e. the .99 quantile of the Gamma distribution line, **b**), at the distance in which 99 per cent of all parcel transactions occur. This distance can be obtained by taking the .99 quantile of the fitted Gamma distribution. With this information, the **b** and **a** lines have been constructed.

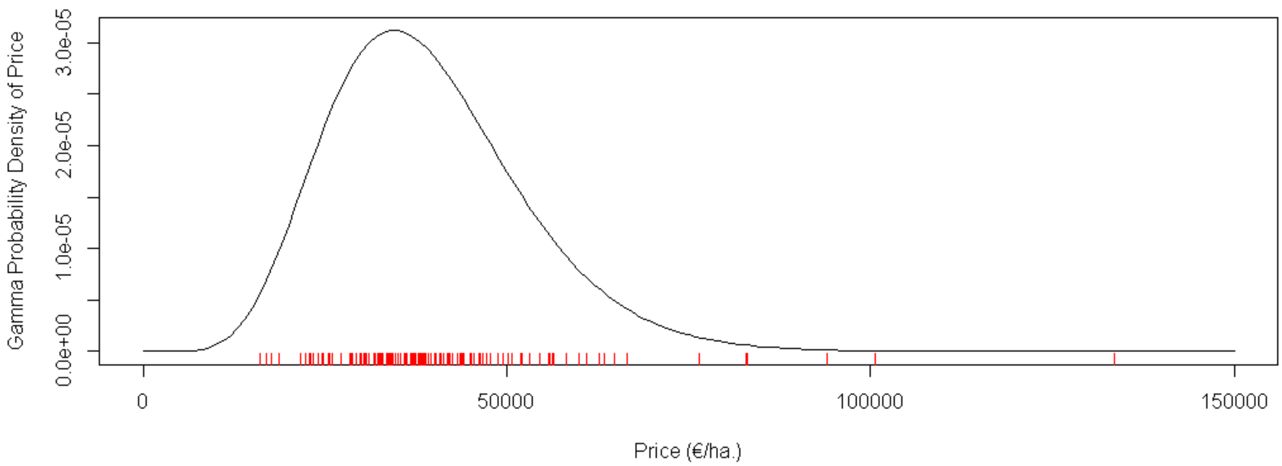


Figure 19 Fitted Gama distribution of price in the first 500 meter (α = 8.44 and γ = 2.16E-4)

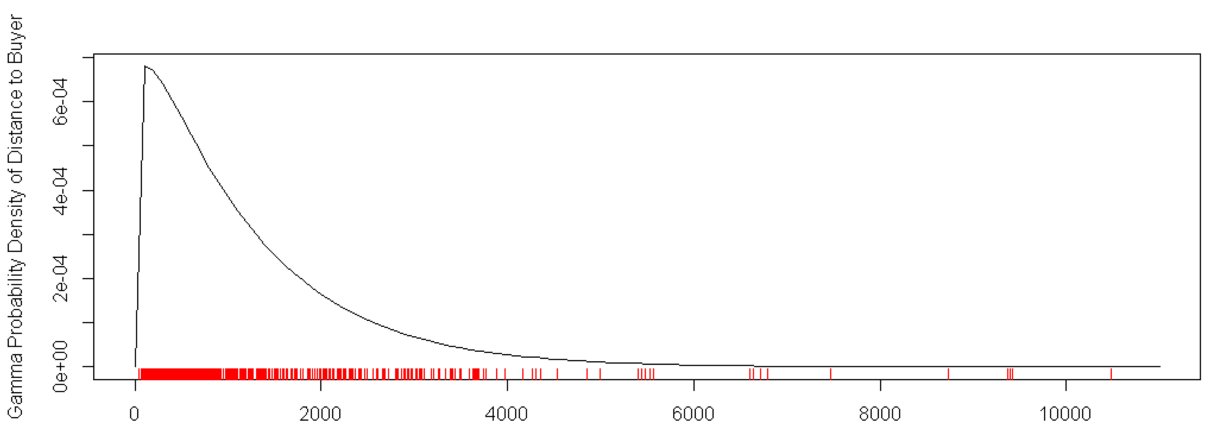


Figure 20 Fitted Gamma distribution of distances of transactions (α =1.11 and γ =0.93E-3)

When fitted to a gamma model, the distance to the buyer in which 99 percent of all transactions take place, sets around 5.5 kilometres (see also Figure 20 at page 35). Additionally, according to the Gamma model of Figure 19 at page 35, 99% of all parcels were sold at a price higher then € 15941 per hectare. Contrarily, 99% of all parcels within 500 meters from the buyer were sold at a price lower than €77095. Within the first 500 meter, also according to Figure 19, the median price (i.e. exactly that price where an equal amount of transactions have a higher and a lower price) sets at € 37614.

The above results in a hypothetical and averaged willingness to pay described by the linear black line *qgamma(.5)* in Figure 21. The willingness to pay declines by about € 12,22 per hectare per additional metre distance to the potential buyer.

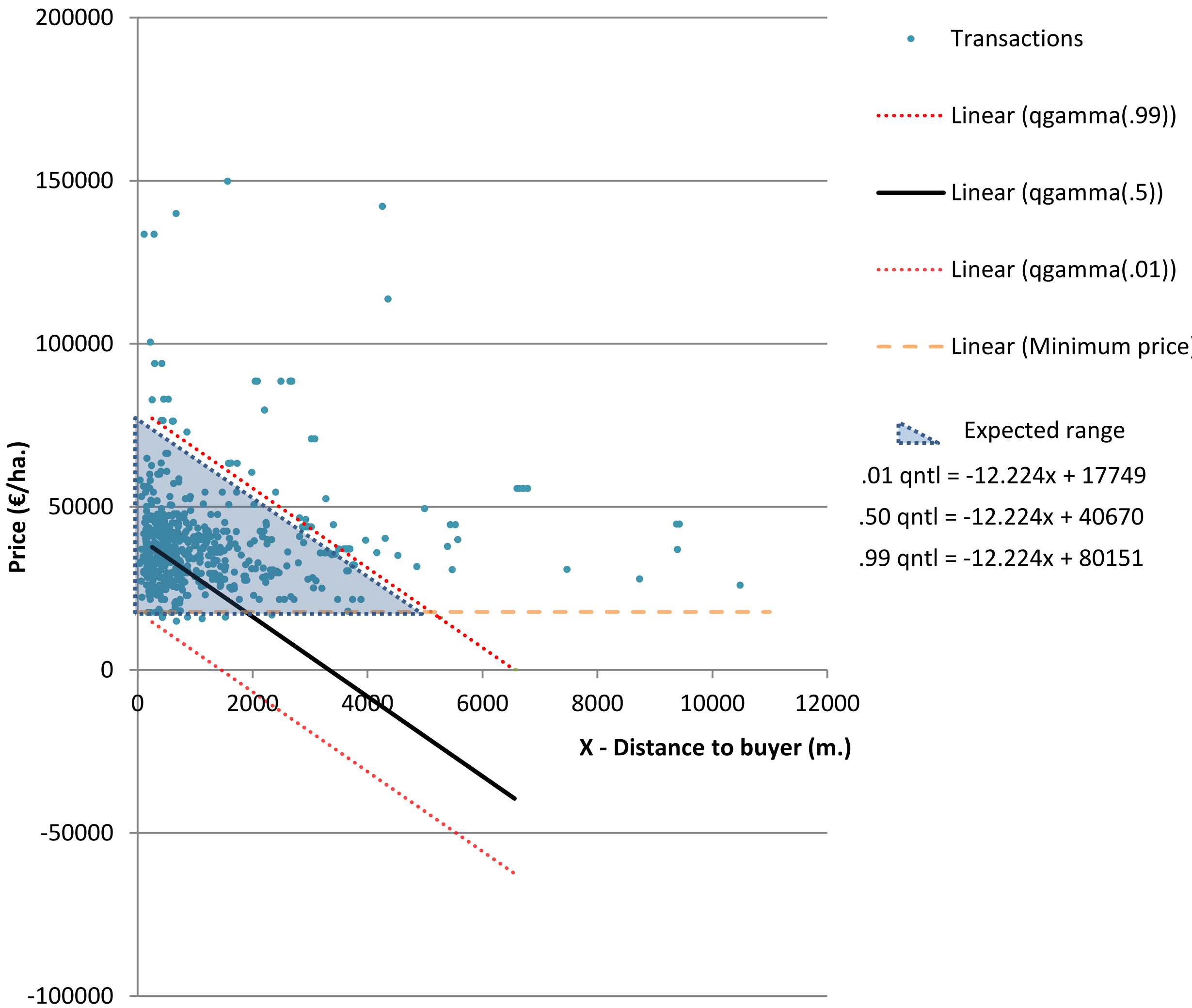


Figure 21 The constructed willingness to pay as a function of distance.

3.2.2 THE HEDONIC PRICE MODEL

Two hedonic price models have been created to investigate if the $WTP^{distance}$ relationship is able to represent actual land prices. The first one, presented inTable 18, explains the actual land transaction prices purely by objective locations aspects. The second one, presented on the right side of Table 18 explains the deviation of the actual transactions price. Only the significant variables are shown. Estimates coefficient values in Table 18 are as expected. Biophysical variables are less important to land transaction prices then the inverse distance to villages and the ecological main structure. The soil texture (as.factor(texture)) is not significant, as it is compared with the null data of soils. For the residual model a similar pattern arises. Additionally, whether an expander currently owns or leases out the parcel matters: a higher price paid for a parcel often results in leasing out a parcel. However, the deviation of the actual land prices to the linear interpretation of willingness to pay by distance is explained worse (R^2 of 0.14 versus 0.20).

Table 18 Hedonic price model coefficient estimates and their probabilities.

Transaction price fit R^2: **0.20**	*Estimate*	*Pr(>|t|)*		*Residual fit* R^2: ***0.14***	*Estimate*	*Pr(>|t|)*	
(Intercept)	1.26E+06	*0.91*			-3.64E+07	*0.04*	*
area	-3.91E+05	0.24			-5.24E+05	0.31	
geom	7.00E+08	0.37			8.53E+08	0.49	
as.factor(texture)1	2.33E+07	0.03	*		4.00E+07	0.02	*
as.factor(texture)2	2.31E+07	0.03	*		3.78E+07	0.02	*
arable	-9.90E+04	0.35			-1.09E+05	0.51	
dairy	2.36E+05	0.05	.		2.90E+05	0.13	
invdistvil	6.18E+09	<2e-16	***		5.75E+09	0.00	***
road_dist	1.08E+02	0.79			3.78E+02	0.55	
invdistehs	-4.47E+08	0.15			-5.30E+08	0.27	
as.factor(ehs_binary)1	-1.83E+06	0.50			-2.03E+06	0.63	
as.factor(property)1	-9.50E+05	0.46			-1.24E+07	0.00	***

3.2.3 LAND EXCHANGE

Two situations of land supply have been modelled. One in which a shrinker supplies his full stock to the market (and is willing to accept any bid higher than his WTA), and one in which he only offers one parcel to the market. If more bids are placed on one parcel, the one that has the highest WTP wins). The number of farms that succeeded in their intensions to exchange land is shown in Table 19. As the supply of land is capped, the number of successful expanders nearly halves. Please note that only 150 parcels are provided, setting the maximum success rate of expanders 52%. Also, the number of land transactions is shown (the spatial distribution of these transactions is shown in Annex I).

Table 19 Number of succeeded land exchanges, and the amount of successful agents.

	Succeeded Land Exchange	
	All parcels	One parcel
Land transactions	936	150
Expanders (287)	200 (70 %)	110 (38 %)(max 52%)
Shrinker (204)	154 (75 %)	150(74%)

Very interestingly is the pattern of farms that did not succeed in their goal to sell land (see next page). For the upper two rendered maps, the colours in Figure 22 indicate the kriged interpolated 'density' of shrinkers. The unsuccessful expanders are mostly situated in low density areas. For the lower two rendered maps, the colours indicate the kriged interpolated density of expanders. Again, the unsuccessful shrinkers mostly group on areas where expanders are less abundant.

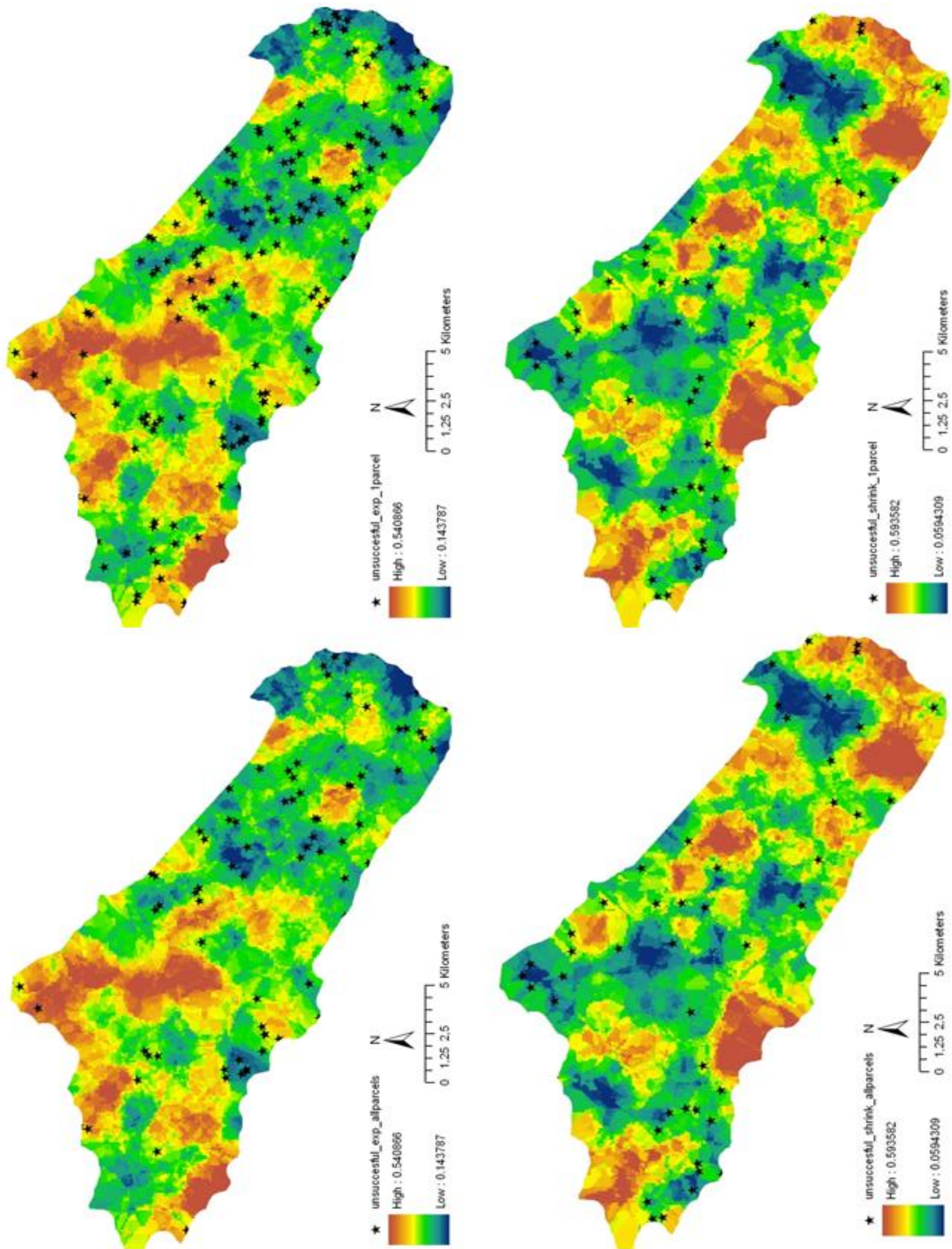


Figure 22 The black stars indicate the farms that did not succeed their land exchange intensions. They are plotted against a background of interpolated probability of their opposite agent types. Clockwise from left under: unsuccessful expanders (all parcels offered), unsuccessful expanders (1 parcel offered), unsuccessful shrinkers (1 parcel offered), unsuccessful expanders (all parcels offered.

4. DISCUSSION

4.1 THE MULTINOMIAL LOGISTIC MODEL

The multinomial logistic model discriminates three farm typologies: expanders, shrinkers and those who remained stable. The explanatory variables involved were the farm area ("area", in hectares), the economic size of the farm ("nge"), the age of the farmer ("age") and whether the farm was dairy based ("type"). Models' fit to the data have been described by measures of goodness-of-fit such as the McFadden R^2 (0.16) and the predictive power (58%). It turned out that the probability for a farm to expand diminishes with age, and increased with economic size and being a dairy farmer. The probability for a farm to shrink increases with age, decreases with economic size and increases with the farm area. The 2009 census data was used to make a map of shrinking, expanding and stable farms, which indicated some spatial clustering of both expanders and shrinkers.

4.1.1 EXPLANATORY VARIABLES

A major pattern in the results of the multinomial logistic model is, among others, the strong dynamics that determine whether or not a farmer is an expander. Dairy farms having a one unit larger "ln(nge)" (comparable to a growth in nge of 54 to 148), have a 5 times higher odds of being an expander. See also Table 13 on page 32. For non-dairy farmers, this is about 2 times as high. This could mean that farmers, and especially dairy farmers cannot intensify their farm without eventually expanding or vice versa. Intensifying before expanding is on par with current findings of Jager and Van Everdingen (2001), who state that for land-based farming, intensifying only holds for the short-term, after which farm expansion takes place. Thus agricultural scaling is an interaction between specialisation, intensification and expansion(Bremmer et al.). A farmer specializes to reduce his fixed cost, as he applies the same land use and management to his total farm. Additionally, the farmer will intensify (not expand, because marginal costs of expanding are higher (van der Heide, Silvis and Heijman 2011)), as after specializing he has more resources (i.e. time, knowledge etc.) to do so. As the farm's maximum efficiency is reached, further income increase can only be achieved by expanding. The demand for land will especially be high for dairy farms, which face, after land-based manure restrictions (Vukina and Wossink 2000), the abolishment of milk quota and therefore are likely to increase their milk production(Jongeneel and Tonini 2009). This means that they will intensify, which eventually increases their demand for land. This cycle can be broken by unfolding non land-based practices, or multifunctional land use, or diversification (Jongeneel, Polman and Slangen 2008). This seems a sustainable way of farming, as demand for multifunctional land use goods are expected to increase in the coming decades (Thijssens 2012). This process will lead to an increase of nge, but without the need of expansion. Multifunctional land use was already existent between 1999 and 2009 (in the Baakse Beek, about 10% of the farms were actively involved in multifunctional land use). Most probably, farms that increased their income only by multifunctional land use were among the stable farm typologies.

Another process of major importance is the positive influence of area to a farms probability of shrinking. The odds of being a shrinker are 3.3 times higher if a farm has one unit of "ln(area)" more land. This is contra-intuitive: as stated above, large farms have low fixed costs and are thus able to easily increase production, resulting in further expanding. Probably, this does not always hold for a (relative to Dutch standards) small-scale agricultural area as The Baakse Beek. Owning many hectares of land may severely reduce farm efficiency as it is labour consuming. In terms of this thesis, this may be expressed in a decrease in the average willingness to accept (WTA) for all parcels by a farm. This is inevitable: theoretically, if a farm owns all 'next-door' parcels, but had acquired some more, these additionally acquired parcels are positioned further away from the farm – thus labeled a lower WTA. Therefore, farms are probably strongly dependent on their spatial configuration of parcels for further expansion or shrinkage. A non-economical spatial distribution of parcels may eventually cause a farm to shrink to increase efficiency. However, a side note must be made. *This effect only occurs if the economic size is not large enough to support the area size.* Would the economic size increase (which always does with area increase), the combined odds of area size increase and nge increase will damp the total odds of becoming a shrinker. Furthermore, as agriculture is in heavy weather, selling land may proof a good strategy to ensure income (if no mortgage is put on the parcel). Note that these are strategies of survival: these farms are by definition large enough to continue farming after shrinkage. These effects could also be caused because there are no quitters contained in the dataset. Because there is no literature explaining these dynamics, another option is discussed: the data may be influenced by sample outliers. Although the data has already been logistically transformed, strongly decreasing the effects of outliers, the model was run again by excluding the 20 largest (in hectares) farms. No significant change in coefficient estimates occurred, thus reducing the probability of large influences by sample outliers.

Additionally, age showed to be a fairly significant explanatory variable in both agent groups. The odds of becoming a shrinker from a one-year increase, is 1.02 versus 1. This means that the chance of becoming a shrinker increases every year with 2%. The odds of becoming an expander decreases every year with 3%. Various reasons may underlie this. For example, old farmers could need money for their retirement and achieve this by selling land. Another survival strategy could be that the farm is too outdated to yield sufficient income. Expanding farms have on average a young farmer at its head. Modern healthy farms are probably better at finding a successor (Potter and Lobley 1992). Furthermore, young famers tend to be better educated than their predecessors which in most cases increases their income (Lockheed, Jamison and Lau 1980). Ultimately inevitable, the age of all farmers will increase every year. This means that a major deal of farms will automatically have an increasing chance of shrinking (if they do not find a successor). Combined with the increased chance of shrinking if the economic size is small, old and small farmers are likely to stop (and thus, in this case, shrinking is not strategy to survive, but a strategy to stop).

The farm area size did not significantly contribute at the $\alpha = 0.05$ level to explanation of variance among the expander type. This can probably be explained by the confounding effect of nge and area. Regardless of the initial area size, for becoming an expander a farmer needs a large economic size. Furthermore, the farm type and the nge*type interaction were not contributing significantly to the model fit of the shrinkers. This implies that there is no higher or lower chance for dairy farms to become a shrinker, nor is it specifically dependent on a dairy farm's economic size. The recipe for shrinking is mainly defined by age, a large area and a small economic size, regardless of the farm type.

Theoretically, would the model be run several times, a farmer can also change his strategy over time. As a farmer expands, he increases his area. A large area is a significant driver for the shrinkers type. However, would he loose land again, former characteristics that previously defined him to be a expander might take over again. As the farmer behaviour becomes complex.

4.1.2 MODEL PERFORMANCE

The performance of the model is assessed here as it provides the basis of the spatial configuration of expanders and shrinkers. Model performance should be treated very delicately. As shown in the "Results" section, the McFadden R^2 yields 0.16. This does not mean 16 per cent of the outcome variance is explained, but that the added value of the regression model is 16%. The model predicts about 60% per cent of the outcome variable correctly, of which 44% is due to chance alone, and 16% is the merit of the regression model. Expander farms were easily distinguished from stable farms, but shrinkers were difficult to distinguish from stable farms. This is a result of the reflections presented at the end of section 2.2.2. In this section it is discussed that the models ability to explain variance of the outcome variable is dependent on the number of agents for each outcome category. This is not discussed in literature, as multinomial logistic models are often based on discrete choice based surveys on behaviour (e.g. (De Groot et al. 2012) or have outcome categories that, contrarily to this study, could not have been transformed from quantitative values to qualitative decision making (e.g. soil categories (Kempen et al. 2009) or land uses (Rounsevell et al. 2006)).

The model struggles to separate shrinking farms from stable farms. This is probably caused by some similarities in characteristics of the shrinking and stable farms. It is expected that a high proportion of stable farms is non-land based as they are not actively taking part in the land markets. At least for a part of the shrinkers' population this hold true too. The odds of a stable farm becoming a shrinker is mostly dependent on area size, and a little on age and nge. In other words, shrinkers are quite similar to stable farms. The additional land they own compared to stable farms is apparently not fundamental to their farm type. Would they be all land-based, they would all reduce their fixed income. Would one draw behaviour stochastically from the farms' behaviour probabilities, the vague distinction between shrinkers and stable farms probably manifests. Expanders are far more distinct from their stable colleagues.

Additionally, in section 2.2.2, it was suggested that the second choice ' mode' has information on agent type probability. In Table 20 it can be seen that this statement holds, as the second choice performs very well where the first choice did not predict correctly. This means that if the residuals of the first choice are predicted with the second choice, 85% of all observations are predicted correctly.

Table 20 Residuals solving with second choice mode

Residual modelling	Residual Second Choice Overall Pred. Power	Residual Second Choice Stable Pred. Power	Residual Second Choice Expanders Pred. Power	Residual Second Choice Shrinkers Pred. Power
	64 %	89 %	59%	50%

However, as it is unknown which part is predicted incorrectly, one can only use the first choice 'mode' . This has serious consequences for the correctness of the number of shrinkers. For the year 1999, the number of shrinkers was severely under-predicted due to deterministically defining behaviour. As the shrinkers were the smallest outcome categorie, they systematically have lower probabilities. This could have been solved by stochastically drawing from the probability distributions of farms. Nonetheless, no complete parcel and farm location data of the year 1999 could be obtained and therefore maps and data of 2009 were used to model land exchange. One might argue that a predictive power of 20 per cent for shrinkers is too little to make assumptions on the consequences of the spatial distribution of those shrinkers. However, note that this small predictive power was caused by under-prediction. In the year 2009, due to developments in nge, age and area, this under-prediction does not occur anymore. We might very well assume that the model is able to predict shrinker better in the situation of 2009 for the year 2010, than for the situation in 1999 for the year 2009.

Table 21 Explanatory variables used by Cotteleer (2008) to predict expanding farms probabilities

Results of Probit Model with Buyer as a Dependent Variable (=1 if Farmer Bought Land in 2003, 0 Otherwise)

Explanatory Variables	Coefficient	*t*-Statistic
Farm is partnership (=1, 0 otherwise)	0.2884***	7.81
Farm is main activity of farm operator (=1, 0 otherwise)	0.0366	1.57
Additional activities are undertaken by farm operator (=1, 0 otherwise)	0.0339	1.49
Age of youngest farm operator (years)	−0.0082***	−11.71
Agricultural land (ha)	0.0028***	11.02
Land used for horticultural open air activities (ha)	0.0073***	6.11
Grassland (ha)	0.0023***	4.43
Fallow land (ha)	0.0490***	4.51
Number of Cattle Size Units (mainly cattle, sheep, goats and horses)	0.0023***	10.31
NGEs (Dutch Size Units), a normative size of the farm activities	0.0001**	2.40
Farmer sold land in the last five years (=1, 0 otherwise)	0.0842***	3.88
Farmer bought land in the last five years (=1, 0 otherwise)	0.2307***	12.26
Constant	−1.6135***	−36.14
Pseudo R^2	0.0497	
Likelihood Ratio Statistic Chi Squared (12)	1541	
Number of observations	85,189	

Addionally, a first step in this research had already been set by Cotteleer (2008). She investigated explanatory variables of land buyers. As can be seen in Table 21 (which is taken directly from her article), most explanatory variables of land buyers are highly significant but explain little of the variance (see pseudo R^2). Apparently, the model presented in this thesis performs better. This is not strange, as intensions of both models differed. This thesis focusses on the prediction of expanders and shrinkers, while the study of Cotteleer had other research ojectives.

The consequences of the results of the multinomial logistic regression can be relevant to our basic understandings of land dynamics. If one assumes that the spatial configuration of shrinking and expanding farms defines the spatial configuration of land scarcity, the socio-economic characteristics of a farm should be acknowledged as the major drivers.

4.2 LAND MARKET DYNAMICS

4.2.1 WILLINGNESS TO PAY

The 'willingness to pay' by distance relationship suggests that the willingness to pay for a hectare of land reduces by about €12,22 for every meter the parcel is farther away from the potential buyer. The relationship determines which farms succeed in their efforts to buy or sell land, and which do not. The assumption that this relationship is linear is hypothetical, but it can also be falsified by creating a model on the basis of the integrals in section 2.4.1. The product of the integrals in section 2.3.1 results in Figure 23. Most strikingly, the shape of the Poisson-distribution in Figure 10 is closely reproduced. However, the amount of transactions falls far behind those of Figure 10. The peak of that curve reaches 60 transactions within distance ranges of 100 metre. This is because that curve represents the whole area, and the curve here only an imaginary circle around an imaginary farm. Furthermore, the peak is situated around 500 metre, here it is situated around 1500 metre. A possible explanation may be that very close to the farm, the WTP – distance relationship should climb a bit more towards a higher WTP. Other explanations could lie in the concept of the willingness to pay (which is further worked out in section 6.1).

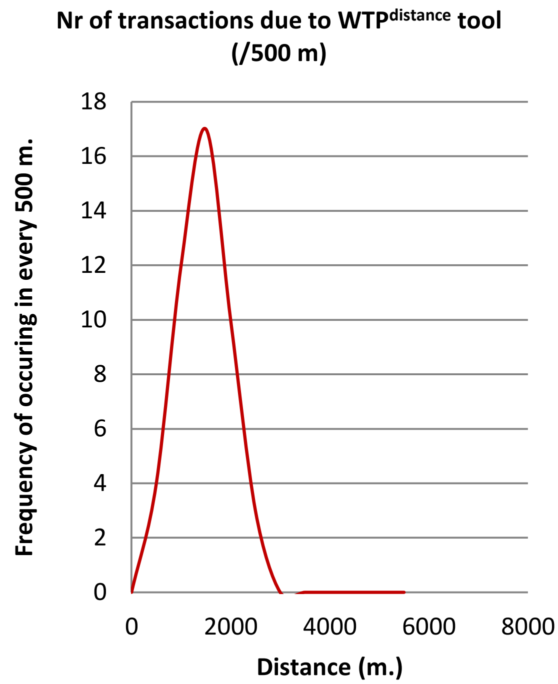


Figure 23 Modeled number of transactions with a linear declining WTP

4.2.2 HEDONIC PRICE MODEL

Two hedonic price models have been created, one for the residuals of the WTP-distance relationship (calculated by comparison of willingness to pay by distance to actual prices paid) and one for the actual prices paid. Would the former model have a better fit (i.e. a higher R^2), the WTP-distance relationship would be a good subjective location proxy for the actual price paid before looking at objective location aspects . This is not the case. This means that the WTP-distance relationship is not able to represent real transaction prices but only gives a measure of influence of distance to a hypothetical willingness to pay. But this should not be taken too serious. First of all, the fits of the hedonic price models were quite low (R^2 0.14 and 0.20) compared to other literature investigating objective location aspects (0.66 for (Garrod and Willis 1992), 0,59 for (Bockstael 1996), although they included a more thorough list of explanatory objective location aspects). This indicates the relative unimportance of objective location aspects in the study area.

Secondly, imagine the situation at the distance wherein about 90 per cent of all farmers buy their land. In the Baakse Beek area, this is around 3500 meters. The WTP-distance relationship would yield a willingness to pay of minus 2000 hypothetical euros. Actual prices at those distances are still around the average of € 40,000. Would an hedonic price model of objective location aspects be fitted, then these objective location aspects supposedly compensate for 42,000 euros. But these objective aspects would compensate every farmer to the same extent, of course.

The relevance of the used reasoning to possible locked land markets and agricultural scaling lies in the concept of objective location aspects. These objective location aspects only determine a minor fraction of the WTP for land. As objective location aspects are of the same (small) importance to every farmer, the spatial configuration of buyers around a parcel determines a major part of the willingness to pay. In addition, the spatial configuration of parcels for sale also determines willingness to pay. It therefore can be assumed that the state of local land markets *is* mostly defined by the spatial configuration of expanders and shrinkers.

4.2.3 LAND EXCHANGE

Based on the situation in which the shrinkers offer all their land for sale, about three quarter of the potential expanders and shrinkers were successful in exchanging land. This number decreased sharply for expanders when the supply of land was reduced to one parcel per shrinker. As shrinkers were almost as successful when supplying only one parcel compared to supplying their full stock of parcels, shrinkers seem to be more dependent on their position than on their supply of land. Expanders and shrinkers that did not succeed within the land exchange model are strongly clustered in areas where their opposites are less abundant.

Before going into the results of the land exchange, some remarks about the land exchange model must be made. The model is an oversimplification of the rural land market. For example, the willingness to pay is only dependent on distance, where income, market power and other factors could be included. Furthermore, there is no minimum willingness to accept. This creates situations where a parcel is sold for prices close to or even under zero. However, the oversimplification is applied to elucidate the possible influence of the distance to the buyer and seller. As unsuccessful expanders or shrinkers appear to be clustered at certain locations, this has worked out well. Now, how could the spatial distributions of potential expander or potential shrinkers have become clustered in the first place? According to the multinomial logistic model, this relies on their explanatory variables. However, except for age, the spatial distribution of area size, economic size and farm type is a product of a long chain of decision making on the land market. Farms that chose to intensify in the past, had eventually no other options but to expand (Jager and Van Everdingen 2001). Probably, a farm can only expand by 'engulfing' his shrinking neighbours. If he has no shrinking neighbours anymore, he is not able to expand any further. Clusters of potential shrinkers are probably formed by another process as shrinking is not very likely a long term farming strategy. It is suggested here that if a shrinker cannot find a market for his land, his neighbour probably cannot either. Thus, the clusters of expander and shrinkers can be seen as 'locked' land markets.

As agricultural scaling can be slowed down by locked land markets, non-land based agricultural sectors (such as poultry and pork) had the opportunity to grow significantly faster in terms of their economic size (and are perceived more as industrial). Moreover, the scarcity of land in The Netherlands probably caused The Netherlands only having a small amount of large farms (see Figure 24). As a reaction, Dutch land-based farms intensified by increasing their production because, as stated before, intensification in The Netherlands is not as expensive as expanding (due to land scarcity). Literature suggests that scaling in agriculture causes homogeneous landscapes by the declining abundance of small landscape properties (such as hedges, unpaved roads, small parcels), and consequently, losses in biodiversity (Tscharntke et al. 2005) and a reduction of the rural liveability(Newby 1979). The findings in this thesis indicate that these processes have been significantly delayed by land scarcity and have not had the same impact as in Spain, the United Kingdom or France. While scaling may have been slowed down by the unfavourable positions of expanders and shrinkers, the resulting intensification still may cause damage to ecosystem services as nutrient fluxes increase(Reinhard, Lovell and Thijssen 1999).

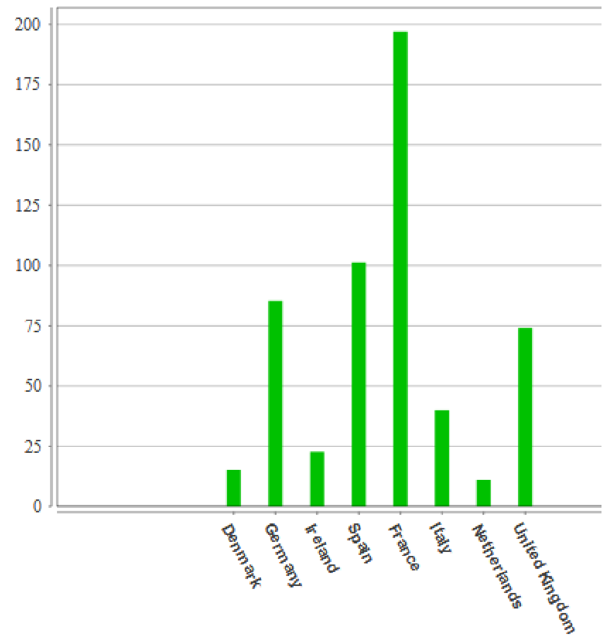


Figure 24 The number of farms having more than 50 hectares for developed countries in the EU (x 1000) (Eurostat 2007a)

5. CONCLUSIONS

Scarcity issues of land can be very local as farms that expand are bound to their location and dependent on the spatial configuration of parcels from farms that shrink. Results of this study show that the spatial configuration of farms can indeed prevent potential expanders from acquiring land or potential shrinkers from emitting land. Compared to other developed countries, this may have slowed down processes of agricultural scaling and its consequences, but increased processes and externalities of intensification. The general objective of this study was to understand the consequences of local land markets by developing a method that captures and explains the spatial distribution of farms that will potentially expand or shrink. A model that predicts farms' behaviour (shrinking or expanding) on land markets was created, followed by a model that simulated land exchange as a function of distance between a parcel and the buyer and seller.

Farm behaviour on the land market can be predicted correctly for 58% by a combination of four socio-economic characteristics. These include the farmer's age, the farm's economic size, the farm's area size and the farm type. Aging reduced the chance of farm expansion and increases the chance of shrinking. Probably because young farmers tend to be better educated, which generally leads to an increase of income. Economically large sized farms have a higher odds of expanding (especially for dairy farms), and lower odds of shrinking. A farm's area size was unimportant to its expanding behaviour, but increased the chance of shrinking as a farm could become inefficient due to its large size.

By constructing a relationship between willingness to pay and the distance between buyer and parcel, parcels that are for sale could be valued by expanding farms. Hypothetically, the willingness to pay declines with €12,22 for each meter distance. The exact drop in willingness to pay by distance could not be confirmed by hedonic price modeling. Nonetheless, relationship provided an insight in the dynamics of local land markets. In a simulation run of farms in 2009, where shrinkers offered only one parcel for sale, only 38 per cent of the expanders succeeded in acquiring land where only four parcels remain unsold. Should shrinkers offer all their land for sale, 70% of all expanders succeeded. Successful shrinkers were more dependent on their spatial configuration than on the number of parcels they offered. Unsuccessful potential expanders that did not succeed, clustered in areas where potential shrinkers were less abundant and vice versa. This is probably caused by historical growth of these clusters, where expanders "eat" their shrinking neighbors until a point they do not have shrinking neighbors anymore.

6. RECOMMENDATIONS

6.1 WILLINGNESS TO PAY

The willingness to pay used in this study is created from a geographers perspective and does not recognize the demand curve used in microeconomic theory(Varian 2006). The demand curve in standard economic theory shows the relationship between price and demanded quantity of a good. As parcels are not bought in ‘quantities’, it is assumed that the WTP curve shows a trade-off between price and demanded quality. In this study, the relationship between the demanded qualities and the WTP are assumed to be linear and equal to all agents. However, Filatova (2009a) suggests to apply the standard demand curve from economic theory to the WTP:

$$WTP = \frac{Y * U^2}{b^2 + U^2}$$

Where:

Y = Income of the agent

U = Parcel’s utility

b = Price of all other goods

Application of this standard demand curve has some major advantages. The utility-bearing characteristics mentioned in section 2.3.4 can be included (i.e. the income of the agent). Increasing Y results in a higher WTP. Likewise, overall economic conjuncture can be included by changing b. Increasing b results in a lower WTP. Increasing U leads to higher WTP. Please recall though, assessing the WTP as a function of distance remains challenging, as there is no information on the effect of distance to the farmers’ utility.

Furthermore the WTP and the bid price could be separated(Filatova et al. 2009a). This allows negotiating power and market power to be introduced. To understand the consequences of this approach, one must first recognize that if neglecting negotiating and market power, transactions prices are set at the point where marginal WTP and WTA of respectively the seller and buyer are equal and are thus only defined by the utility of the parcel to the seller and buyer. However sellers and buyers try to maximize returns from their transactions.

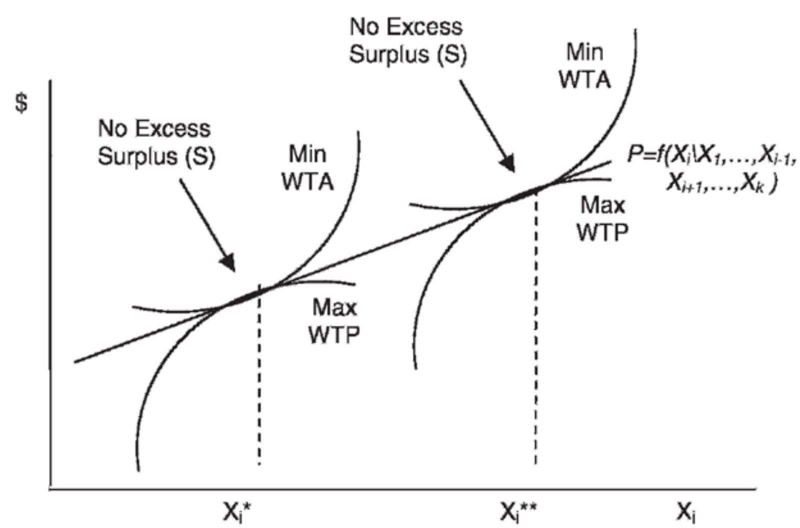


Figure 25 Hedonic Price Model standard situation without negotiating or market power. P is the hedonic price function of characteristic *i*. At some levels (X_i^* and X_i^{**}) of characteristic *i*, slopes of WTA and WTP are equal, and therefore the maximum WTP and minimum WTA are equal(Cotteleer, Gardebroek and Luijt 2008).

This means that buyers likely present a bid price lower than their WTP (and sellers ask more than their WTA), leaving room for negotiation (e.g. based on the personal characteristics of the buyers and sellers (Cotteleer et al. 2008)). Additionally, the state of local land markets may shift bid prices from the WTP. If the local land market has more sellers than buyers, it is a buyers' market – causing ask and bid prices to shift downwards from respectively WTA and WTP. If there are more buyers than sellers, it is a sellers' market, causing ask and bid prices to shift upwards from respectively the WTA and the WTP.

These sophisticated bid price adjustments are analysed by Cotteleer et al. (2008) and (LeBaron, Tesfatsion and Judd 2006). However they have only been included in a quite theoretical agent based model on urban fringe land markets (Filatova et al. 2009b) which does not recognize *local* land markets by a $WTP^{distance}$ tool. As the concept of willingness to pay is fundamental to the model presented in this study, it will be very interesting to implement the bid price adjustments.

6.1 AGENT BASED MODELLING

Current agent based models could benefit from the approach used in this study. First it is discussed how agents are generally defined and studied. Please keep in mind how *this* study is conducted and which difficulties normally encountered in studying farm typologies with an agent based approach, have been worked around unnoticed. After the section on 'agent based modeling', the application of the land exchange model created for this thesis to land use and cover change is discussed. In the field of land use and cover change (LUCC), this model of land exchange may prove to be a very welcome tool.

Contemporary research has used agent typologies to simplify reality (Barnes et al. 2011). In general it can be said that an agent of land imposes a use to the land he or she currently owns, this as a result of external (e.g. price) and internal(e.g. age, education, values, soil type) incentives, resulting in a specific land cover. Completely understanding this process would require an exact understanding of the agent's multidimensional situation. Typologies of agents are made to simplify reality and can be constructed in different ways. Currently, two schools dominate agent typology research, one using a cultural approach and one using a behavioural approach.

The 'cultural approach' in typology research advocates qualitative methodologies. It relies heavily on understanding language, meaning, respresentation, identity and difference. It studies social relationships, how people share norms, values and views(Gylfason et al. 1999) and how this can be observed in specific farming typologies(Lowe et al. 2002). These outcomes are, ultimately, based on in-depth interviews and sociological expert knowledge. Representativeness is very unclear in this approach. As the method is precise, but not accurate, it may be impossible to empirically base an ABM with this approach(Janssen and Ostrom 2006). An example of research on farming typologies is (Roep, van der Ploeg and Leeuwis 1991), which is a cultural study towards sustainability of farming in The Achterhoek region.

Behavioural approaches are quantitative methodologies that focus "on the motives, values and attitudes that determine the decision-making processes of individuals" (Fishbein and Ajzen 1975), with individuals being farmers. Behavioural approaches in agricultural studies have their roots in the missing 'satisficing' concept(Simon 1957, Burton 2004). Men do not necessarily take economically optimal decisions, but also try to optimize social, psychological or other goals. A very clear example is a case study called 'the Edinburgh study of decision making on farms' by Willock (1999). Willock designed scales of attitudes, objectives and implementation on certain decision-making, and collected her information via a social survey. It gives some information about farmers' attitudes, but it is not very objective as actual decision-making cannot be witnessed. Notwithstanding deficits in accurateness and preciseness, this approach is often used to empirically inform an ABM in decision-making science. An example of the application of this approach in an ABM is (Karali, Rounsevell and Doherty 2011). In this study, typologies are extracted from social surveys using factor and cluster analysis. These are statistical tools that group respondents into certain typologies when their answers are more or less the same (but still their actual decision-making can be very heterogeneous). In accordance with Robinson (2007) (saying a variety of data gathering approaches yields the most reliable result), Karali et al. also used in-depth interviews to base their findings.

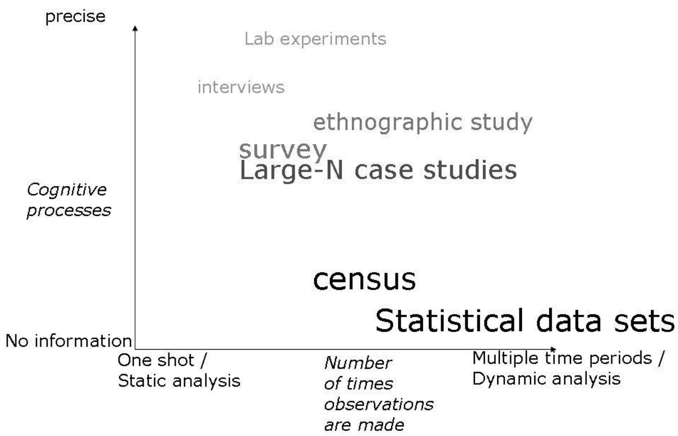


Figure 26 Methods to empirically inform ABM's. The horizontal axis denotes the temporal character; the vertical axis indicates the insights given in decision making. There is also a third 'axis'. The darker the text, the more respondents are involved. From: (Janssen and Ostrom 2006).

What strikes most in these approaches of typology research, is the fact that neither of them is able to significantly represent real decision making. Firstly, the cultural approach lacks representativeness and quantified data, the behavioural approach lacks preciseness and observations of actual decision making. Agent typologies should be made from objective data which represents behaviour c.q. decision making, on a sample large enough to be representative. This is problematic, as on-farm decision making is inherently a qualitative process, which can only be observed on-site and real-time (and therefore, no literature could be found that combined the qualitative characteristics decision making with a very large sample). Should one observe behaviour quantitatively, one needs two 'slices' in time of the agents studied. Within this timespan, decision making must have been observed. This study has dodged these difficulties and used databases of two years of the variables of interest.

6.2 LAND USE AND COVER CHANGE

Despitefully, economists had their hands full on land prices analysis and could not foresee their research into land markets possibly provides a basis for a new agent based model of LUCC. Although this thesis focusses on processes of land market dynamics, the methods presented here can be very relevant for LUCC modelling with agent based models. A framework that addresses human characteristics and their dynamics with the environment is needed to study LUCC (Rounsevell et al. 2012). It is thought that agent based modelling is a good candidate in explaining underlying processes of LUCC, due to the inclusion of individual decision-makers (i.e. agents) (Parker et al. 2003, Bousquet and Le Page 2004).

Agricultural scaling is probably one of the most prominent processes in land use and cover change. In developed countries with mostly specialized farms*, decisions on categorically changing of farm type* (and thus land use) are probably not likely to occur. A farms combination of high degree of specialism, high debts and high age reduce decision flexibility. LUCC is therefore not to arise from on-farm land use decision making by a single farmer, but from a parcel of land changing owner. For example, LUCC occurs when a parcel of land is bought from a dairy farmer by an arable farmer. The land use will change from dairy (i.e. most probably a grass – maize rotation), to arable (e.g. a grain-beet-onion rotation). Therefore, transactions of land are most probably a main driver of LUCC. Imagine the situation that on a sample of about 500 farms, 50 % of all dairy farmers (which is 50% of 50% of the sample, thus 25% in total) change their land use (e.g. to pig breeding or arable farming) in a time span of 10 years. Considering the investments (i.e. money, knowledge, mind set) involved, this does not seem very realistic. When reading the following arguments, the conclusion must be that on-farm LUCC decision making is not a major dynamic in Dutch LUCC:

- just like some other EU countries, farms in The Netherlands are (highly) specialized (see Figure 27). For the Netherlands this also holds that these farms are highly intensive(Eurostat 2007b). For their specialism and intensiveness, these farms have invested considerable in (human) capital, farm structure and farm strategy and are therefore not very flexible to change land use;
- agricultural holdings in The Netherlands are often heavily financed (Meulen 2008). It is not strange to state that low solvability reduces flexibility on farm decision making;
- socio-psychological research by Lokhorst (2011) suggests that the farm type is a major part of a farmers identity – obvious to those that have spoken to farmers(Schnabel 2005, Roep et al. 1991).

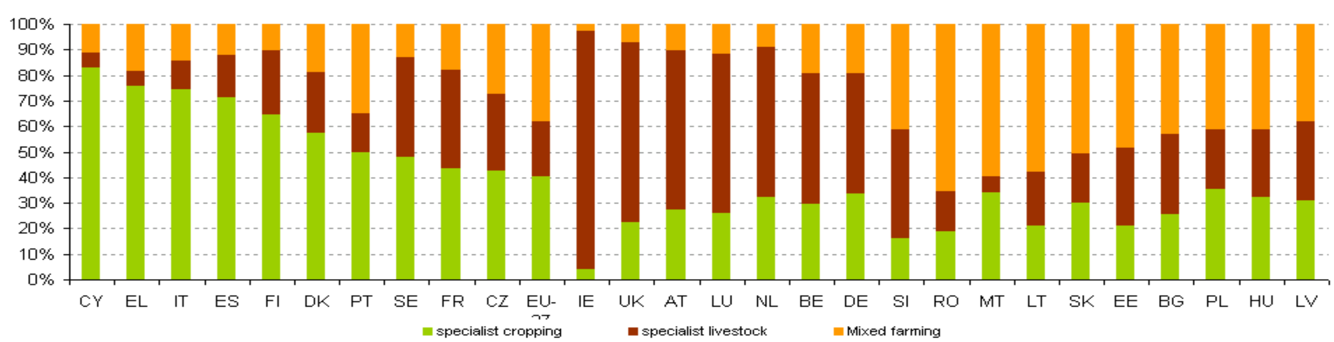


Figure 27 Share of specialized agriculture for each of the 27 EU member states(Eurostat 2007b). More than 90% of the farms in The Netherlands are specialized either in livestock or in cropping.

If it is assumed that the former idea of on-farm decision making on LUCC does not hold anymore, how can agents (i.e. autonomous farms) still affect land use? Of course, agents still apply their land use to their fields. A dairy farmer in The Netherlands will cover its fields with a grass-maize rotation, an arable farmer may cover his fields with a grain-beets-onions rotation and a pig farmer may hold some arable land. But, as that does not change, how *does* land use change? This can only be achieved if a farm's parcel spatial configuration of parcels changes (i.e. agricultural land transactions). This economical concept requires an agent that sells his or her parcel, and an agent which is willing to buy that parcel. *The change in ownership results in LUCC.*

In summary, explaining agricultural scaling by processes of shrinking and expansion of farms is very different from current approaches in agent based modeling of LUCC:

- in current approaches agents do not really represent land use decision-makers as no actual agent decision-making is involved. For example, they are based on socio-economical characteristics (Bakker and van Doorn 2009), or only based on subjective (to the researchers) management styles (Valbuena, Verburg and Bregt 2008, Karali et al. 2011).
- In many studies, the spatial extent of each agent has not made it to agent based models. Agents have mostly been represented by cellular automata(e.g. see (Robinson and Brown 2009)), resulting in an unlikely high clustering of land use (as in reality, farms have a scattered parcel distribution). In the method presented in this study, each farm has a unique spatial parcel distribution.
- Processes of on-farm decision making are inherently qualitative, making approximations by quantitative ABM models difficult (Rounsevell et al. 2012). By approaching decision making by interactions in local land markets, a quantitative (economic) tool is used to make qualitative decisions (a land transaction).
- In this new framework agents make decisions not based on intentions , but based on actual observations of the particular behaviour of the agent (i.e. farm expansion or shrinkage).
- While ABM has loads of potential, its data needs may require

Short History of LUCC Modelling

In the starting days of modeling LUCC in The Netherlands (in the early nineties), two schools of Linear Programming (LP) modelling emerged. One for regional LUCC (Rabbinge and Van Latesteijn 1992, de Wit et al. 1988, van Ittersum et al. 2004), and one that modelled a single farm(Berentsen and Giesen 1995, Wossink 1993). LP is mechanistic: by programming a list of linear income and cost functions, as well as preconditions and constraints, a Pareto optimum is calculated. However, LP stood far from practice, as Wossink (1993) also noted in his conclusions of his PhD. thesis.

To make a spatial explicit study on land use, Stoorvogel (2004) designed the Trade-Off Analysis (TOA) tool. Economic and bio-physical variables are measured on a sample of fields, indicating the trade-off between these variables based on the 'opportunity costs' principle(Varian 2006). TOA is very helpful as a (policy) tool, quickly showing how trade-offs are made given certain conditions – especially when no databases are available yet. However, as TOA emphasizes more the consequences of LUCC, it does not model spatial LUCC.

The CLUE-S model (Veldkamp and Fresco 1996, Verburg et al. 2002) calculates the spatial distribution of LUCC. More or less similar to regional LP, it uses a regional set of services that must be met by the land use given the local biophysical situation (i.e. land suitability, infrastructure). LUCC only occurs if the regional services are still met and the new land use has a clear value increase. The approach is based on empirical relationships between land use and its drivers. While CLUE-S makes a decision for every raster cell, it is not based on actual decision-making (cells do not decide of course). Agent based modelling is therefore the logical next step to analyse LUCC. However, it still needs a lot development.

sample surveys, in-depth interviews, geo-referenced data, remote sensing data and so on (Robinson et al. 2007). The presented framework uses existing farm census and cadastre data.

ACKNOWLEDGEMENTS

As it has always seemed a bit forced to me to thank those whom eventually will never read this thesis, I firstly want to thank my main supervisor Martha Bakker. She read everything. Much of the work presented in this thesis is the fruit of endless discussions with Martha on challenges I faced. Martha let me think I was in control of my research, but in the end she was, of course. I thank her for the time spent on my research and the fact she was always available for discussion (and when she was not, she fixed someone to do it for her). We went on a business-trip to Edinburgh to paint the city r.. to have great discussions at the University of Edinburgh on the creation of a new agent based model. I learned a lot about scientific projects, scientific discussions and the scientific scene. Furthermore, Martha hates exclamation marks written by lads!

Then I would like to thank those who probably will never read my thesis but who contributed greatly. Tom Kuhlman, probably the man acknowledged most in the entire world, provided me with loads of (meta)data on farms and farmer behaviour. I would like to thank Derek Robinson for commenting my early proposal drafts, Mark Rounsevell for taking Martha and me out to diner to the Scotch Single Malt Whisky Society in Edinburgh (although the CARE project paid the evening and I had not even written down one letter of this thesis). As a second supervisor, I would like to thank Jetse Stoorvogel. He also arranged things when Martha could not supervise me. I also have to thank Jan Klink (a former roommate, now working at the Ministry of ELI, who will not have expected this), as he provided me with a report about the relevancy of land exchange to agricultural scaling. It steered me in the right direction.

I thank those farmers, eventually the direct objects of my thesis, for their time and their lovely coffee. Also, Noud Janssen from LTO and LLTB and Ruud Tijssens of Agrifirm provided me with lots of new information on the agricultural sector. The Land Dynamics group is thanked for their funding, as well as the CARE project. Anne Marike Lokhorst of the Communications Sciences Group is thanked for her time in the preliminary stage of this study, as well as Dirk Roep from the Rural Sociology Group. Additionally, I thank Laurens Van Buuren and Diederick Koetsier for commenting on some drafts. Edwin Haasjes I thank for providing me with his car to visit farmers and CEOs. Now probably Kas Lange is already asking why he is not mentioned yet, as I am mentioned in his thesis. Now you are in mine.

Last but not least I thank Floor, but I guess she prefers chocolate over being acknowledged.

BIBLIOGRAPHY

Alonso, W. 1964. *Location and land use: toward a general theory of land rent*. Harvard University Press.

Armsworth, P. R., G. C. Daily, P. Kareiva & J. N. Sanchirico (2006) Land market feedbacks can undermine biodiversity conservation. *Proceedings of the National Academy of Sciences,* 103, 5403-5408.

Bakker, M. M., M. Sonneveld & B. J. Brookhuis (In Prep.) Soil as a spatial determinant of land use in time.

Bakker, M. M. & A. M. van Doorn (2009) Farmer-specific relationships between land use change and landscape factors: Introducing agents in empirical land use modelling. *Land Use Policy,* 26, 809-817.

Barnes, A. P., J. Willock, L. Toma & C. Hall (2011) Utilising a farmer typology to understand farmer behaviour towards water quality management: Nitrate Vulnerable Zones in Scotland. *Journal of Environmental Planning and Management,* 54, 477-494.

Berentsen, P. B. M. & G. W. J. Giesen (1995) An environmental-economic model at farm level to analyse institutional and technical change in dairy farming. *Agricultural Systems,* 49, 153-175.

Bockstael, N. E. (1996) Modeling Economics and Ecology: The Importance of a Spatial Perspective. *American Journal of Agricultural Economics,* 78, 1168-1180.

Bousquet, F. & C. Le Page (2004) Multi-agent simulations and ecosystem management: a review. *Ecological Modelling,* 176, 313-332.

Bremmer, J., A. G. J. M. Oude Lansink, K. D. Olson, W. H. M. Baltussen & R. B. M. Huirne. Analysis of Farm Development in Dutch Agriculture and Horticulture.

Burton, R. J. F. (2004) Reconceptualising the ‘behavioural approach’ in agricultural studies: a socio-psychological perspective. *Journal of Rural Studies,* 20, 359-371.

Cavailhès, J. & P. Wavresky (2003) Urban influences on periurban farmland prices. *European Review of Agricultural Economics,* 30, 333-357.

CBS (2009) NEG Typering 2009.

---. 2010. Land- en tuinbouwcijfers 2010.

Chavas, J.-P. & A. Thomas (1999) A Dynamic Analysis of Land Prices. *American Journal of Agricultural Economics,* 81, 772-784.

Cho, S.-H. & D. H. Newman (2005) Spatial analysis of rural land development. *Forest Policy and Economics,* 7, 732-744.

Cotteleer, G., C. Gardebroek & J. Luijt (2008) Market Power in a GIS-Based Hedonic Price Model of Local Farmland Markets. *Land Economics,* 84, 573-592.

Croissant, Y. 2010. Estimation of multinomial logit models in R: the mlogit packages. Université de la Réunion.

De Groot, C., F. B. C. Daalhuizen, F. van Dam & C. H. Mulder (2012) Once an outsider, always an outsider? The accessibility of the Dutch rural housing market among locals and non-locals. *Journal of Rural Studies,* 28, 302-313.

de Wit, C. T., H. van Keulen, N. G. Seligman & I. Spharim (1988) Application of interactive multiple goal programming techniques for analysis and planning of regional agricultural development. *Agricultural Systems,* 26, 211-230.

Eurostat. 2007a. Farms larger than 50 ha.

---. 2007b. Specialized And Mixed Farming.

Filatova, T., D. Parker & A. Van der Veen (2009a) Agent-Based Urban Land Markets: Agent's Pricing Behavior, Land Prices and Urban Land Use Change. *Journal of Artificial Societies and Social Simulation,* 12.

Filatova, T., D. Parker & A. van der Veen (2009b) Agent-based urban land markets: Agent's pricing behavior, land prices and urban land use change. *JASSS,* 12.

Fishbein, M. & I. Ajzen. 1975. *Belief, attitude, intention, and behavior: an introduction to theory and research*. Addison-Wesley Pub. Co.

Garrod, G. D. & K. G. Willis (1992) Valuing goods' characteristics: An application of the hedonic price method to environmental attributes. *Journal of Environmental Management,* 34, 59-76.

Gerbens-Leenes, P. W., S. Nonhebel & W. P. M. F. Ivens (2002) A method to determine land requirements relating to food consumption patterns. *Agriculture, Ecosystems & Environment,* 90, 47-58.

Gylfason, T., T. T. Herbertsson & G. Zoega (1999) A MIXED BLESSING. *Macroeconomic Dynamics,* 3, 204-225.
Healy, R. G. & J. L. Short (1979) Rural Land: Market Trends and Planning Implications (Journal of the American Planning Association). *Journal of the American Planning Association,* 45, 305-317.
Hosmer, D. W. & S. Lemeshow. 2000. *Applied Logistic Regression*. Wiley.
In Lee, K. & J. J. Koval (1997) Determination of the best significance level in forward stepwise logistic regression (Communications in Statistics - Simulation and Computation). *Communications in Statistics - Simulation and Computation,* 26, 559-575.
Irwin, E. G. & N. E. Bockstael (2002) Interacting agents, spatial externalities and the evolution of residential land use patterns. *Journal of Economic Geography,* 2, 31-54.
Jager, W. H. & J. H. Van Everdingen. 2001. Kosten- en opbrengstenvoordeel melkveebedrijven door schaaleffecten. In *Agri-Monitor*. LEI.
Janssen, M. A. & E. Ostrom (2006) Empirically Based, Agent Based Models. *Ecology and Society,* 11, art 37.
Jongeneel, R. A., N. B. P. Polman & L. H. G. Slangen (2008) Why are Dutch farmers going multifunctional? *Land Use Policy,* 25, 81-94.
Jongeneel, R. A. & A. Tonini. 2009. Dairy Quota and Farm Structural Change: A Case Study on the Netherlands. EAAE.
Karali, E., M. D. A. Rounsevell & R. Doherty (2011) Integrating the diversity of farmers' decisions into studies of rural land-use change (Earth System Science 2010: Global Change, Climate and People). *Procedia Environmental Sciences,* 6, 136-145.
Kempen, B., D. J. Brus, G. B. M. Heuvelink & J. J. Stoorvogel (2009) Updating the 1:50,000 Dutch soil map using legacy soil data: A multinomial logistic regression approach. *Geoderma,* 151, 311-326.
King, D. A. & J. A. Sinden (1994) Price Formation in Farm Land Markets. *Land Economics,* 70, 38-52.
LeBaron, B., L. Tesfatsion & K. L. Judd. 2006. Chapter 24 Agent-based Computational Finance. In *Handbook of Computational Economics*, 1187-1233. Elsevier.
LEI. 1999. BINTERNET.
Lockheed, M. E., T. Jamison & L. J. Lau (1980) Farmer Education and Farm Efficiency: A Survey. *Economic Development and Cultural Change,* 29, 37-76.
Lokhorst, A. M., H. Staats, J. van Dijk, E. van Dijk & G. de Snoo (2011) What's in it for Me? Motivational Differences between Farmers' Subsidised and Non-Subsidised Conservation Practices. *Applied Psychology,* 60, 337-353.
Lopez, R. A., F. A. Shah & M. A. Altobello (1994) Amenity Benefits and the Optimal Allocation of Land. *Land Economics,* 70, 53-62.
Lowe, P., H. Buller & N. Ward (2002) Setting the next agenda? British and French approaches to the second pillar of the Common Agricultural Policy. *Journal of Rural Studies,* 18, 1-17.
Luijt, J. (1987) An attempt to explain changes in regional agricultural land prices in the Netherlands. *Tijdschrift voor sociaalwetenschappelijk onderzoek van de landbouw,* 2, 189-201.
Luijt, J. & M. J. Voskuilen. 2011. Grond voor schaalvergroting. 37. LEI.
Maddison, D. (2000) A hedonic analysis of agricultural land prices in England and Wales. *European Review of Agricultural Economics,* 27, 519-532.
Metz, W. C., M. Morey & J. Lowry. 1990. Hedonic Price Theory: Concept and Applications. In *High Level Radio Active Waste Management*.
Meulen, H. A. B. v. d. (2008) Vreemd vermogen per land- en tuibouwbedrijf sterk toegenomen (in Dutch) Debts per farm strongly increased. *Agri-monitor,* 2008 / 14, 7-8.
Mühlenberg, M. & J. Slowik. 1997. *Kulturlandschaft als Lebensraum*. Quelle & Meyer.
Newby, H. 1979. *Green and pleasant land?: social change in rural England*. Hutchinson of London.
Parker, D. C., S. M. Manson, M. A. Janssen, M. J. Hoffmann & P. Deadman (2003) Multi-Agent Systems for the Simulation of Land-Use and Land-Cover Change: A Review. *Annals of the Association of American Geographers,* 93, 314-337.
Potter, C. & M. Lobley (1992) AGEING AND SUCCESSION ON FAMILY FARMS: The Impact on Decision-making and Land Use. *Sociologia Ruralis,* 32, 317-334.
Rabbinge, R. & H. C. Van Latesteijn (1992) Long-term options for land use in the European community. *Agricultural Systems,* 40, 195-210.
Reinhard, S., C. A. K. Lovell & G. Thijssen (1999) Econometric Estimation of Technical and Environmental Efficiency: An Application to Dutch Dairy Farms. *American Journal of Agricultural Economics,* 81, 44-60.

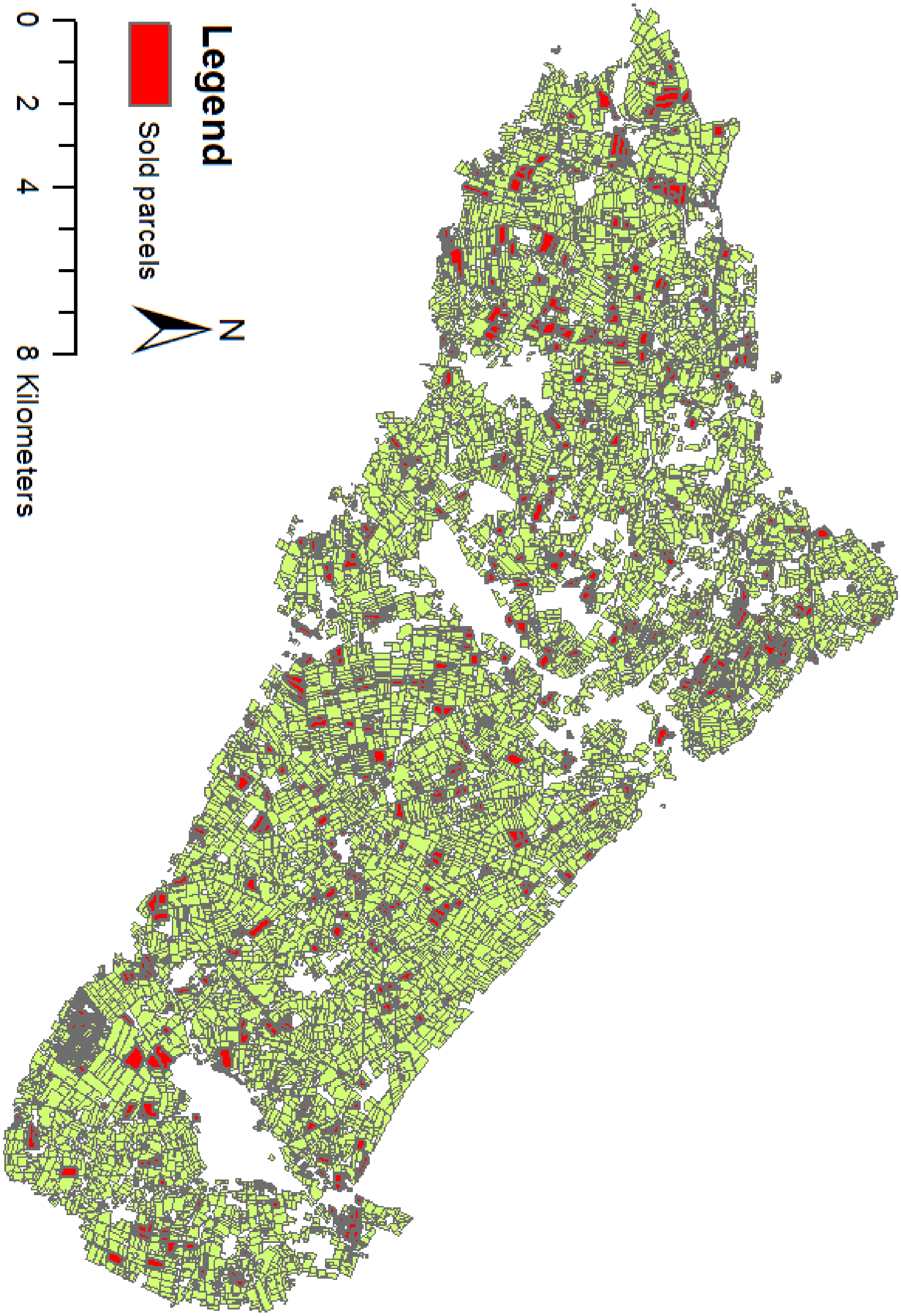
Robinson, D. T. & D. G. Brown (2009) Evaluating the effects of land-use development policies on ex-urban forest cover: An integrated agent-based GIS approach. *International Journal of Geographical Information Science,* 23, 1211-1232.
Robinson, D. T., D. G. Brown, D. C. Parker, P. Schreinemachers, M. A. Janssen, M. Huigen, H. Wittmer, N. Gotts, P. Promburom, E. Irwin, T. Berger, F. Gatzweiler & C. Barnaud (2007) Comparison of empirical methods for building agent-based models in land use science (Journal of Land Use Science). *Journal of Land Use Science,* 2, 31-55.
Roep, D., J. D. van der Ploeg & C. Leeuwis. 1991. *Zicht op duurzaamheid en kontinuiteit : bedrijfsstijlen in de Achterhoek*. Wageningen: Landbouwuniversiteit.
Rosen, S. (1974) Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy,* 82, 34-55.
Rounsevell, M. D. A., I. Reginster, M. B. Araújo, T. R. Carter, N. Dendoncker, F. Ewert, J. I. House, S. Kankaanpää, R. Leemans, M. J. Metzger, C. Schmit, P. Smith & G. Tuck (2006) A coherent set of future land use change scenarios for Europe (Scenario-Based Studies of Future Land Use in Europe). *Agriculture, Ecosystems & Environment,* 114, 57-68.
Rounsevell, M. D. A., D. T. Robinson & D. Murray-Rust (2012) From actors to agents in socio-ecological systems models. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 367, 259-269.
Schnabel, P. 2005. Waarom blijven boeren? (in Dutch) Why do farmers stay? Sociaal en Cultureel Planbureau.
Silvis, H., A. Oskam & G. Meester. 2008. *EU-beleid voor landbouw, voedsel en groen: van politiek naar praktijk*. Wageningen Academic Publishers.
Simon, H. A. 1957. *Models of man: social and rational; mathematical essays on rational human behavior in society setting*. Wiley.
Smith, V. K., C. Poulos & H. Kim (2002) Treating open space as an urban amenity. *Resource and Energy Economics,* 24, 107-129.
Stoorvogel, J. J., J. M. Antle, C. C. Crissman & W. Bowen (2004) The tradeoff analysis model: integrated bio-physical and economic modeling of agricultural production systems. *Agricultural Systems,* 80, 43-66.
Thijssens, R. 2012. Agrifirm, Director Corporate Affairs.
Tscharntke, T., A. M. Klein, A. Kruess, I. Steffan-Dewenter & C. Thies (2005) Landscape perspectives on agricultural intensification and biodiversity – ecosystem service management. *Ecology Letters,* 8, 857-874.
Valbuena, D., P. H. Verburg & A. K. Bregt (2008) A method to define a typology for agent-based analysis in regional land-use research. *Agriculture, Ecosystems & Environment,* 128, 27-36.
Van Bruchem, C. & H. Silvis. 2008. Agrarische structuur, trends en beleid. LEI.
van der Heide, M., H. Silvis & W. J. M. Heijman (2011) Agriculture In The Netherlands: Its Recent Past, Current State And Perspectives. *APSTRACT: Applied Studies in Agribusiness and Commerce,* 5.
van der Ploeg, B. 2001. *Het weigevoel in het groene hart van de randstad: een studie onder melkveehouders in het westelijk veenweidegebeid naar hun bereidheid en mogelijkheden zich te ontwikkelen van productieboer tot plattelandsondernemer*. Wageningen Universiteit.
van Ittersum, M. K., R. P. Roetter, H. van Keulen, N. de Ridder, C. T. Hoanh, A. G. Laborte, P. K. Aggarwal, A. B. Ismail & A. Tawang (2004) A systems network (SysNet) approach for interactively evaluating strategic land use options at sub-national scale in South and South-east Asia. *Land Use Policy,* 21, 101-113.
Varian, H. R. 2006. *Intermediate microeconomics: a modern approach*.
Veldkamp, A. & L. O. Fresco (1996) CLUE-CR: An integrated multi-scale model to simulate land use change scenarios in Costa Rica. *Ecological Modelling,* 91, 231-248.
Venables, W. N. & B. D. Ripley. 2002. Modern Applied Statistics with S. . Springer.
Verburg, P. H., W. Soepboer, A. Veldkamp, R. Limpiada, V. Espaldon & S. S. A. Mastura (2002) Modeling the Spatial Dynamics of Regional Land Use: The CLUE-S Model. *Environmental Management,* 30, 391-405.
Vorage, R. 2010. Investeringsagenda agrarische sector Baakse Beek – Veengoot. LTO Noord, Waterschap Rijn en IJssel.
Voskuilen, M. J. & C. M. van Elk. 1990. *Motieven van glastuinders voor vestiging in de regio Aalsmeer*. Den Haag: L.E.I.
Vukina, T. & A. Wossink (2000) Environmental Policies and Agricultural Land Values: Evidence from the Dutch Nutrient Quota System. *Land Economics,* 76, 413-429.
Willock, J., I. J. Deary, M. M. McGregor, A. Sutherland, G. Edwards-Jones, O. Morgan, B. Dent, R. Grieve, G. Gibson & E. Austin (1999) Farmers' Attitudes, Objectives, Behaviors, and Personality

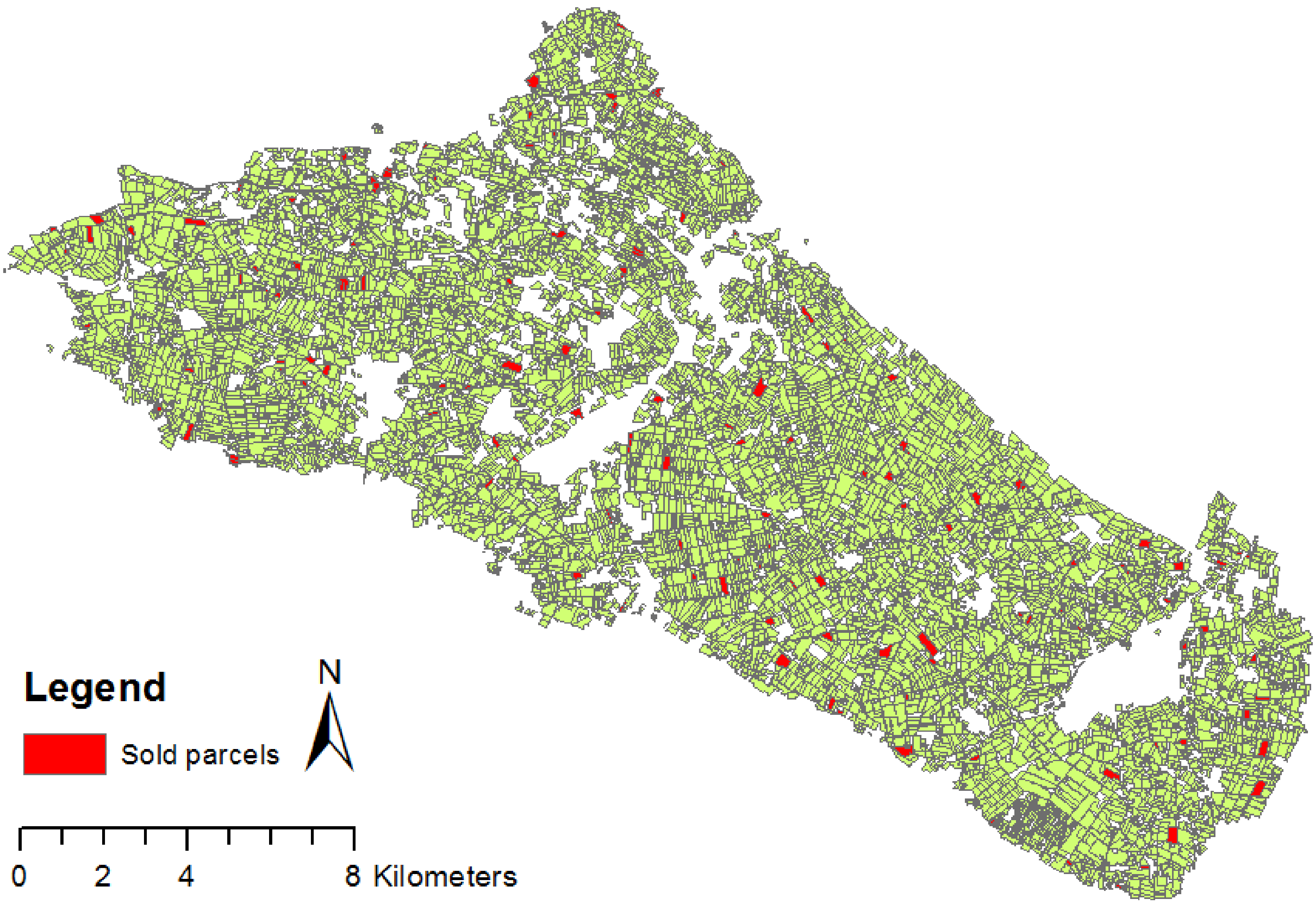
Traits: The Edinburgh Study of Decision Making on Farms. *Journal of Vocational Behavior,* 54, 5-36.
Wossink, A. 1993. Analysis of future agricultural change : a farm economics approach applied to Dutch arable farming. [S.l.]: Wossink.

LIST OF FIGURES

LIST OF TABLES

ANNEX I




Legend
Sold parcels
N
0
2
4
8 Kilometers