sustainability

MDPI

Article

Agricultural Land vs. Urbanisation in Chosen Polish Metropolitan Areas: A Spatial Analysis Based on Regression Trees

Wojciech Sroka [1,*], Jaroslaw Mikolajczyk [2], Tomasz Wojewodzic [1] and Boguslawa Kwoczynska [3]

1 Institute of Economic and Social Sciences, University of Agriculture in Krakow, Al. Mickiewicza 21, 31-120 Krakow, Poland; rrtwojew@cyf-kr.edu.pl
2 Institute of Administration and Economics, State Higher Vocational School in Tarnow, Mickiewicza 8, 33-100 Tarnow, Poland; rrjmikol@cyf-kr.edu.pl
3 Department of Agricultural Land Surveying, Cadaster and Photogrammetry, University of Agriculture in Krakow, Al. Mickiewicza 24/28, 31-120 Krakow, Poland; rmkwoczy@cyf-kr.edu.pl
* Correspondence: w.sroka@ur.krakow.pl; Tel.: +48-502-408-636

Received: 26 February 2018; Accepted: 15 March 2018; Published: 16 March 2018

Abstract: The goal of this paper is to explore intra-regional differences in factors determining land use. We built spatial regression tree models to assess the factors determining the share of agricultural area in municipalities of selected Polish metropolitan areas in 2010. The analyses are static, with the value of exogenous variables presented as an average for the longest possible period preceding the year 2010. We analysed the impact of socio-economic processes, natural conditions, and farming characteristics on the share of agricultural land in the surface area of particular municipalities in metropolitan areas. Based on the concept of economic rents that says that the way land is used is determined by economic rent, we have shown that the most important factor with an impact on the share of agricultural land is the number of enterprises per 10,000 people of working age. Other very important factors have been found to be the quality of environmental conditions of agricultural production, population density, and net migration. It was noted that with an increase in the rate of enterprises, as well as an increase in population density and net migration, the share of agricultural land falls, and a high quality of agricultural production comes with a relatively high share of agricultural land in the surface area of the municipalities analysed.

Keywords: agricultural land; competition for land; economic rents; land use factors; Polish metropolitan area

1. Introduction

A new global phase of development in urbanisation, i.e., metropolisation, has taken on particular importance in recent years. This is one of the most important processes responsible for a functional, as well as a social and an economic change, in settlement systems [1,2]. Metropolisation leads to a concentration of specialised, unique, and rare functions with worldwide and transregional reach in selected cities. It is also a cause of transformation in the economy, in terms of the ways land space is used, as well as in society and in culture [3,4]. In practice, economic data shows that agriculture and agricultural land present in metropolitan areas are particularly exposed to marginalisation and weakening of their productive, social, and environmental functions [5,6]. One of the most important negative consequences of the process of metropolisation is the consumption of land, in particular the loss of high-quality agricultural land [7–9]. The disappearance of arable land is especially pronounced in developing countries [10,11], but is also occurring in Europe, as well. This is confirmed by research conducted as part of the project "Peri-urban Land Use Relationships—Strategies and Sustainability Assessment Tools for Urban-Rural Linkages (PLUREL)" [2]. It has been indicated that the pressure of urbanisation causes both significant changes in available space and the degradation of metropolitan areas (MAs). It has been emphasised that environmentally attractive areas, as well agricultural and open areas, are developed for the construction of housing, communication infrastructure, and business activity in the broad sense [2,9]. Research done in Poland also underlines the fact that the contemporary agricultural land in metropolitan areas is treated as a reserve for other more profitable activities [12,13].

In Poland, the process of shaping the structure of land ownership started anew after 1990, with dynamic changes taking place in land use [14]. It was especially visible in metropolitan areas, where competition for land was the highest, and the development of private sector enterprises and growing urbanisation resulted from conversion of agricultural land [15]. Parcels of farmland, especially those situated within the administrative borders of MA cores (and other cities), have received only very weak legal protection, so the scale of land conversion to non-agricultural purposes was mainly determined by market mechanisms [13,16,17]. This results in a very large differentiation in the share of agricultural land within the surface area of municipalities. In the Polish metropolitan areas selected for analysis, around 50% of land, on average, is used for agricultural purposes, with the percentage varying by an average of 20 percentage points. Research conducted in other countries shows that the way land is used within metropolitan areas is mainly determined by socio-economic factors, including growing urbanisation [8,10]. From an economic perspective, urban expansion is the result of market forces. As long as the marginal benefit of urban land use is greater than that of agricultural land use, urban expansion occurs [18]. This paper formulates a thesis that economic rents (presented in Section 2.1) play a decisive role in determining the relation of the share of farmland to the share of land used for non-agricultural purposes.

The main aim of the paper was to assess factors with an impact on the differentiation in the share of agricultural land in the overall surface area of municipalities of selected Polish metropolitan areas.

2. Theoretical Background

2.1. Theorising Land Use as an Economic Framework

In economic theory, land is regarded as a special asset providing space for locating economic activities, infrastructure, and dwellings, as well as amenity services and aesthetic value [19,20]. In the academic literature, the issues of land change and land cover constitute an area of central interest mainly for urban planners and architects [21–23], but they are also increasingly examined in the field of economics and economic geography. Economists, especially classical economists, were at first unable to indicate the causes of spatial differentiation of economic activity, with Starrett [24] even presenting the Spatial Impossibility Theorem. Later however, as the classical school of location developed, the problem of the anti-spatiality of economics was, at least partly, solved [25]. In order to give at least some idea of the complexity of the issues of land use change, as well as competition for land, we will discuss a few concepts based on the notion of economic rents. According to the authors of this paper, an economic rent is of key importance in explaining the differentiation of land cover in metropolitan areas.

At least since the formulation of bid rent theory by Alonso [26], land use changes and land cover have been usually explained in the economic sciences in terms of the maximisation of utility by a landowner [27]. The foundations for bid rent theory were laid by Ricardo and von Thünen, who developed a land rent theory. In the literature of mainstream economics, Ricardo is considered the author of the land rent theory of differential rent [28]. However, Czyżewski and Matuszczak [29] point out that other precursors can be considered to include Smith and Malthus, who differentiated four forms of land rent—differences in fertility, location, extra capital expenditures, and general rent. Having said that, from the point of view of the development of spatial economics, the greatest contribution was made by the works by von Thünen [25], who noted that the use of space was determined by rent, i.e., income from agricultural production reduced by production costs. He suggested that, where the natural features of the farm, such as climate, soil, topography, and other factors remain constant, the location of agricultural production was determined by the distance or costs of transportation to the urban market. With increasing distance from the town, the land will progressively be given up to products that are cheap to transport in relation to their value [30]. Though this theory is obsolete, its main assumptions about the very large role of location as a factor in land use changes should not be underestimated [31]. The direct approach applies the Thünian model of rural land-use allocation to the problem of urban-rural land conversion [32]. This was also explicitly articulated by Sinclair [33], who argued that, with urban sprawl, increasing competition for land comes from non-agricultural uses. He noted that, in many advanced industrialised parts of the world, the basic forces determining agricultural land use near urban areas are associated with urban expansion. Where these forces are in operation, the agricultural pattern quite often is one of increasing intensity with distance from the city. Non-agricultural uses which bring higher rents "push" agricultural production beyond cities [33].

An urban location model linked to von Thünen's theory has been developed by Alonso [26]. His model can be regarded as the basis for household location choice [20]. Alonso's bid-rent theory explained the relationship between land prices and land use as follows: in a competitive land market, land-users seek to maximise their utility, land being purchased/rented by the bidder offering the highest bid, i.e., the potential land-user able to derive the highest rent from land. Therefore, land is expected to be used for the purpose which brings the greatest utility, taking into account the relative benefits of alternative land uses [19]. Alonso's model [26] inspired numerous researchers and, currently, at least a dozen or so different models are used to present the problem of competition for urban land [34]. For instance, Konagaya [35] presented the so-called Generalised Thünen Models (GTMs), which, in the conditions prevalent in Asia, provided a very good explanation of the phenomenon of competition for space among different uses. The GTMs also reveal the relation of land use with strong rent-bidding power. Urbanisation driven by globalisation is the prime factor producing land-use changes. Examining three means of land use, Konagaya [35] noted that the land closest to the centre was allocated for urban uses, with agricultural production located a little further away, and forest land still further away. Thus, the margin of transference from urban land uses determines the beginning of agricultural land uses. The author observed that as a city grows (e.g., measured by the size of population), these zones shift, with agricultural land use and forest land use being pushed outside.

In summary, it should be stressed that, in most theories of production location, the key role is played by the distance from the markets, i.e., essentially city centres (also cores of metropolitan areas). Differences arising from the location of specific activities in relation to the MA cores are to a large extent a result of their utility/profitability. Activities that require relatively low labour and capital input give way to more intensive, and usually more profitable, ones (principle of utility maximisation). Close to the city centre, land prices are very high (very high competition for land), therefore, only the most profitable activities generate profits/income high enough to cover the costs of engaged production factors (including mainly the land). As noted by Wästfelt and Zhang [30], as well as Mazzocchi et al. [9], agriculture, which is a low-profitability sector, usually loses in the competition for land. The zones of the location of different land uses, as presented by the different authors, are flexible and change as a city develops. Uses bringing the lowest economic rent are pushed beyond the areas subject to urbanisation. When new enterprises (and jobs) are created, and the population grows as a result (with growth in housing construction as a consequence), the agricultural zone slowly decreases in size and starts to be used for non-agricultural purposes.

The relationships presented here are based mainly on the assumptions of neoclassical economics without taking into account public intervention. In reality, the operation of the mechanism of economic rents is disrupted, mainly by administrative decisions or the adoption of local development plans. For that reason, the shares of farmland and other areas in cities will not always be determined by the mechanisms presented. This is of great importance, especially in countries where an active spatial planning policy is pursued.

2.2. Factors of Land Use with Special Reference to Farmland

When analysing drivers determining the share of agricultural land in metropolitan areas, both internal and external ones are indicated as important [9]. Drivers are usually grouped into natural constraints (geophysical factors), socio-economic factors, the institutional framework, and reasons related to unadapted agricultural systems, including, in particular, land fragmentation [8,10,36]. In our opinion, drivers determining land use can be divided into at least four groups: socio-economic factors, which show the impact of processes of urbanisation, environmental, and fixed geographic features, as well as reasons related to the features of agricultural systems and the institutional framework (Table 1).

Table 1. Factors affecting land use/land-use change in metropolitan areas.

Category of Factors	Potential Variables	Effects on Share of Agricultural Land in Overall Surface Area
Socio-economic	– population density, – net migration rate, – economic and employment growth, – labour characteristics,	Increase in population density, migrations, and an increased number of enterprises/decreasing rate of unemployment have an impact on growing demand for land for non-agricultural purposes, resulting in farmland being converted to non-agricultural uses [8,10,37,38].
Fixed geographic features	– distance to the city centre, ports, roads, – road networks, – availability of infrastructure, – quality of environmental conditions (including soil quality, water conditions, climate, slope, elevation),	The closer to the city centre, ports and roads, the bigger the competition for land and the stronger the pressure on its conversion to non-agricultural uses [10,36]. Higher productivity of agricultural land results in higher economic rent from every unit of surface area. Users give up agricultural production and convert land to non-agricultural uses less often [8,19,39].
Features of agricultural systems	– profitability of farms, – structure of farms (share of small farms, land fragmentation), – full time farmers and having a successor,	The higher the profitability of agriculture, the less often farmland is converted to non-agricultural uses. Advantageous structures (large farms, full time farmers, etc.) foster higher profits. [9,16,30,39–41].
Institutional framework	– protected areas, – spatial planning, – fees for agricultural land conversion,	Policy for the protection of prime farmland could be effective in saving agricultural land from development for non-agricultural uses [7,23,42].

Source: own elaboration.

The first group, i.e., socio-economic factors, refers mainly to the influence of processes of urbanisation on land cover. In metropolitan areas, in accordance with the theories presented in Section 2.1 (maximisation of utility by the landowner), urban pressure encourages land conversion from agricultural to urban uses, while the structural weakness of agriculture makes conversion easy [9]. To live and work, humans need space, so population growth, as well as high population density [8], migration [37], economic and employment growth [38] lead to urban land development. In accordance with the models developed by Alonso [26] and Konagaya [35], urban land use brings higher economic rents than agricultural land use, therefore, areas (e.g., municipalities) with high population density, a large number of non-agricultural enterprises, etc., have a relatively low share of agricultural land.

The second group of factors refers to environmental and fixed geographic features. Land cover will be determined by the distance of the land from the city centre, ports, road networks [10,36], and the quality of environmental conditions of agricultural production, including the quality of land, water conditions, climate [8,39]. The distance of a land parcel to markets (von Thünian ideas) and its 'quality' of land, in terms of geo-physical characteristics (Ricardian ideas), determine the land-use decision of each landowner [39]. High competition for land parcels in "good" locations (near the market, at ports, by rivers, etc.) contributes to the increase in their prices. Consequently, in accordance with the principle of utility maximisation, they will be allocated for uses bringing high economic rents. This is because high costs of the purchase or leasing of land have to be offset by high income.

Features of agricultural systems are another important factor determining land cover. Weakness of agriculture resulting from low profitability of farms and a flawed structure of farms (prevalence of small farms and land fragmentation), as pointed out, among others, by Mazzocchi et al. [9] and Xie et al. [40], leads to land being, first, abandoned, and then converted to non-agricultural uses. Grădinaru et al. [41], as well as Hagedorn [16], stress that this problem concerns Poland in particular, which suffers from significant fragmentation of farms. Additionally, Wästfelt and Zhang [30] note that with high land prices in metropolitan areas, farmers are unable to increase either the area or intensity of production, and derive relatively lower economic rents as a result. Even Sinclair [33] stressed that the greater the chances are of urban land uses taking over, the lesser the chances of maintaining agricultural production or increasing its intensity will be, especially in the long run.

Factors relating to the institutional framework are most difficult to assess, but they play a great role in preserving farmland in metropolitan areas [23]. Measures taken by public institutions, e.g., through various forms of protection of green areas [42] and appropriate provisions made in development plans (e.g., not permitting abandonment of land of very good quality), may limit the spatial expansion of cities, correcting the operation of market mechanisms, i.e., the effect of economic rents [8].

3. Materials and Methods

3.1. Study Area

Poland is a country located in Central and Eastern Europe with a relatively low degree of urbanisation. Due to the historical background, i.e., a long period of the loss of statehood, wars, etc., industrialisation and urbanisation processes which started relatively late, i.e., around two centuries later than in most countries of Western Europe [43], Poland is still a country with a relatively significant spatial importance of agriculture, as farmland accounts for over 60% of its surface area, with around 0.48 ha of agricultural land per inhabitant.

The analyses in this paper cover six metropolitan areas—one large one: the Warsaw metropolitan area, four medium-sized ones: the Krakow, Wroclaw, Poznan, and Tricity metropolitan areas, and one small one: the Lublin metropolitan area (Figure 1). They were selected in such a way as to reflect the large diversity of environmental and economic conditions in different parts of Poland. The delimitation of these areas was based on development documents and strategies adopted by regional authorities [44–49]. In total, 280 municipalities were selected for analysis, including six cities constituting the cores of selected (monocentric) metropolitan areas and 274 municipalities from surrounding metropolitan areas (outer zone).

In total, the analysed metropolitan areas (MA) cover around 9.6% of the surface area of the country and are inhabited by 9.4 million people in total, i.e., around 24% of Poland's population (Table 2). By far the largest metropolitan area has formed around Warsaw, and is inhabited by over 3.2 million people. The Krakow, Tricity, and Poznan metropolitan areas have similar populations (around 1.5 million inhabitants each), while the Lublin metropolitan area, inhabited by around 609,000 people, is definitely the smallest. The average population density in the cores of metropolitan areas is 2440 people/km^2, while in outer zone it is around 162 people/km^2, with the national average being 123 people/km^2. Taking into account areas outside cities and towns, the average population density in Poland is 53 people/km^2.

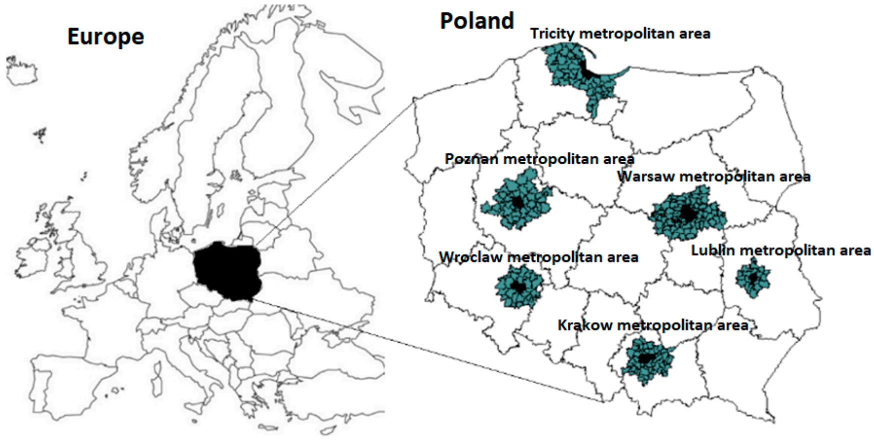

Figure 1. Metropolitan areas which qualified for the study. Source: own elaboration based on [44–49].

Table 2. General characteristics of the analysed metropolitan areas (2016).

Metropolitan Areas	Number of Municipalities	Number of People (in Thousand)		Population Density (People/km^2)		Share of Overall Surface Area in Relation to the Country's Surface Area (%)
		Core	Outer Zone	Core	Outer Zone	
Warsaw MA	81	1754.0	1451.3	3391	217	2.3
Krakow MA	51	765.3	752.7	2341	201	1.3
Tricity MA	55	747.6	816.2	1804	129	2.2
Wroclaw MA	27	637.7	403.4	2177	115	1.2
Poznan MA	45	540.4	880.3	2063	148	2.0
Lublin MA	21	340.5	268.3	2308	136	0.7
Total	280	4785.4	4572.1	2440.7	162	9.6

Source: own elaboration based on data from Poland's Central Statistical Office (GUS).

Proceeding to the main research problem, it should be highlighted that, in 2010, agricultural land of individual farms accounted for around 49.9% of the surface area in the Polish metropolitan areas analysed (Table 3). However, in the cores of metropolitan areas, this share varied from 20% in the Tricity to 48.6% in Lublin, which shows a very large differentiation.

Table 3. Selected characteristics of land use in Polish metropolitan areas in 2010.

In Details:	Percentage of the Agricultural Land of Farms in the Total Area in 2010 (%)			Standard Deviation (SD)
	Mean	Min.	Max	
Core	27.8	20.0	48.6	0.11
Outer zone				
– including urban municipalities *	29.8	0.0	67.6	0.21
– including municipalities directly bordering on the MA core	48.2	2.5	77.5	0.18
– other municipalities located in outer zone	52.2	20.3	94.8	0.16
Metropolitan areas in total	49.9	0.0	94.8	0.22

* In Poland municipalities are classed as urban (consisting of a town or city), urban-rural (consisting of a town together with its surrounding villages and countryside), or rural (not containing a town). Source: own elaboration based on data from Poland's Central Statistical Office (GUS).

In outer zone municipalities, this differentiation is even greater and varies from 0% in the municipality Hel (Tricity metropolitan area) to almost 95% in the municipality Domaniów (Wroclaw metropolitan area). The general trend is that, in urban municipalities, the share of agricultural land in the overall surface area is smaller compared to other units. However, in some cities, agricultural land predominates over other land uses (the share of agricultural land may even reach 67.6% of the overall surface area). Similarly, in municipalities directly bordering on the cores, where urbanisation pressure is very strong, the percentage of agricultural land exceeds 50% of the overall surface area in almost half the municipalities, and 48% of land on average is used for agricultural purposes. Thus, the question that should be raised is what local conditions or processes have led to such a significant differentiation.

3.2. Data Collection and Definition of Research Variables

The main sources of data included results of the 2010 Agricultural Census, information available in the Local Data Bank of the Main Statistical Office of Poland and the academic literature. Analyses concerned differentiation in the share of agricultural land held by individual farms in the overall surface area of the analysed municipalities of metropolitan areas. Areas belonging to legal forms of organisation other than individual farms were intentionally omitted, as the data collection methods used by the Polish Central Statistical Office (GUS) do not make it possible to identify the actual location of the land of farms and enterprises of legal persons. This is because, in some cases, the seat of a farm/agricultural enterprise is located in a metropolitan area, whereas most of its land is outside of it. As pointed out in a publication by the GUS [50], this problem does not apply practically to individual farms.

In this paper, the endogenous variable is the share of agricultural land held by individual farms in the overall surface area of municipalities in 2010 (Table 4). This variable has a normal distribution, therefore, an arithmetic average was used in the analyses. The analyses are static in character, with the value of exogenous variables presented as an average for the longest possible period preceding the year 2010. It should be stressed at this point that the share of agricultural land in the surface area of the analysed municipalities is a result of socio-economic changes that occurred in the past, among other things. The list of variables explaining the differentiation in the share of agricultural land in the overall surface area of the municipalities was compiled based on both theoretical notions and empirical studies conducted by other authors. A potentially broad range of variables grouped in three sections: socio-economic factors, environmental and fixed geographic features, and the structure of agriculture, was qualified for analysis. They are directly related to the discussion presented in Section 2.2.

Table 4. A set of predictors used in the process of modelling.

Factors	Variable Name	Definition of Variable and Unit	Range of Variation/Size
		Endogenous variable	
	Agricultural area	Share of agricultural land held by farms in the overall surface area of a municipality—in 2010 (%)	AV = 49.9 Min = 0 Max = 94.8
		Exogenous variables	
Socio-economic	Type of municipality	Type of municipality—as of 2010 (dummy variable urban, urban-rural, rural)	Urban municipality—44 Urban-rural municipality—75 Rural municipality—161
	Population density	Population density of a municipality—average for the period 1995–2010; (people/km^2)	AV = 367.8 Min = 12.3 Max = 4027.4
	Net migration rate	Internal net migration per 1000 people—average for 2003–2010 (person)	AV = 5.3 Min = −8.8 Max = 37.3
	Commuting to work	Number of people coming in to work per 1 person going out to work—2006 (person)	AV = 0.71 Min = 0.05 Max = 13.7

Table 4. *Cont.*

Factors	Variable Name	Definition of Variable and Unit	Range of Variation/Size
Socio-economic	Unemployment rate	Share of registered unemployed people in the population of working age—average for 2003–2010 (%)	AV = 7.3 Min = 2.5 Max= 19.2
	Entrepreneurship	Economic entities including natural persons-average for 2002–2010 (entities/10,000 population of working age)	AV = 1449.7 Min = 631.9 Max = 5475.4
	Large enterprises	Economic entities employing over 50 people—average for 2002–2010 (entities/10,000 population of working age)	AV = 10.1 Min = 0 Max = 47.4
Environmental and fixed geographic features	Agricultural Production Space Valuation Ratio—APSVR	Agricultural Production Space Valuation Ratio (point) *	AV = 66.6 Min = 36.2 Max = 108.6
	Protected natural areas	Percentage of land under various forms of environmental protection in the overall surface area of municipality—average for 1996–2010 (%)	AV = 26.7 Min = 0.0 Max = 72.9
	Public roads	Length of municipality hard surfaced roads per 100 km^2 of surface area—average for the period 2003–2007 (km/100 km^2)	AV = 66.7 Min = 8.2 Max = 662.5
	Spatial planning	Share of surface area of municipalities covered by spatial development plans—average for 2009–2010 (%)	AV = 46.5 Min = 0.0 Max = 100.0
	Distance to the core	Distance from the centre of the municipality to the centre of the MA core (km) **	AV = 36.6 Min = 0 Max = 79.1
	Travel time to the core	Travel time from the municipality's centre to the centre of MA core (minutes) **	AV = 43.8 Min = 0 Max = 85
	Surface area of municipality	Surface area of a municipality—average for 1996–2010 (thou. ha)	AV = 10.7 Min = 0.5 Max = 51.7
	Direct border with the cores	Including municipalities that share a border with the MA core (dummy variable: yes/no/core)	Yes—75 municipalities No—199 municipalities Core—6 municipalities (cities)
Structure of agriculture	Share of small farms	Share of farms with an area less than 5 ha in the overall number of farms—average for 1996, 2002 and 2010 (%)	AV = 60.6 Min = 15.7 Max = 98.8
	Share of economically strong farms	Share of farms with an economic size above 15,000 Euro Standard Output—2010 (%)	AV = 16.9 Min = 0.0 Max = 59.5

* Agricultural Production Space Valuation Ratio is a synthetic measure that takes into account the quality of: soil, climate, topography, and water conditions. Its maximum value is 125 points. The average value for Poland calculated in this way was 66.6 points [51]. ** Distance and travel time according to Google Maps assuming that the shortest route by car is set. Source: own elaboration based on data from Poland's Central Statistical Office (GUS).

The paper includes the most important exogenous variables available in the Polish system of public statistics at the level of municipalities (the objects of the studies). The applied method of regression trees did not require that variables have a normal distribution, nor was it necessary to verify whether correlation existed between them. The individual variables were characterised by a relatively high volatility.

The analyses omitted external institutional factors, as they are similar across all of the examined municipalities, which are subject to the same law. Only information concerning the share of plots of land covered by spatial development plans in relation to the overall surface area of the municipalities studied was taken into account in the model. Such plans specify the location of agricultural land, hence, it can be assumed that the fact of adopting such a plan may have an influence on the share of land for agricultural use in relation to overall surface area. Nevertheless, in the municipalities studied, still only about 46.5% of the land is covered by spatial planning, and the majority of scholars [16,17,52] note that the farmland listed in local development plans in Poland receives very weak legal protection, and the spatial policy pursued in Poland does not limit the operation of market mechanisms, allowing for land speculations and uncontrolled growth of urbanised areas.

3.3. Data Analysis: Regressions Trees

Regression trees are classified as exploratory methods of data analysis. The paper uses one of the most advanced methods of building regression trees, which is based on the CART (Classification and Regression Trees) algorithm. Regression trees have a number of advantages over other models, including highly popular regression models. Their structure is non-parametric, small trees are readily interpretable, there is no global sensitivity to outliers, and they are able to handle non-linear relationships well. Further advantages are that regression trees are suitable for a larger number of variables, thus avoiding the necessity of a strong pre-selection of the variables. The problem of collinear variables is automatically handled by the stepwise procedure of the approach. Hence, regression trees select and combine the most relevant factors in a well-defined model structure [53]. However, regression trees have some limitations and weaknesses. As regression trees use the "divide and conquer" method, they tend to perform only a few highly-relevant attributes [54]. Small changes in data can alter a tree's appearance drastically and thereby alter the interpretation of the tree if not managed with caution [55]. These weaknesses were eliminated using the interactive regression trees technique. The CART methodology could be criticized because it does not provide a statistical output, such as a confidence interval, by which to quantify or support the validity of the findings [55].

The aim of the tool propagated by Breiman et al. [56] is to search for a set of logical splitting conditions of the "if—then" type. During the construction of a tree, recursive splits of a set of observations into disjoint sub-sets are conducted [57]. Using a variance-minimising algorithm, regression tree models repeatedly partition the data to determine increasingly homogenous sub-groups, based on partition criteria of the independent variable. The objective of regression tree analysis is to derive a structure, according to the independent variable, which produces the most homogenous nodes. Dependent variable data is split into a series of left and right child nodes derived from the primary nodes [58].

In choosing the best splitter, the program seeks to maximise the average "purity" of the two child nodes. The "impurity" of a node t of the tree is defined as the sum of squared deviations $\boldsymbol{i}(\boldsymbol{t}) = \sum(\boldsymbol{y} - \overline{\boldsymbol{y}}_t)^2$, where $\overline{\boldsymbol{y}}_t$ is the sample mean of response variable Y in t and the sum is over the y values in t. The split of t into child nodes t_L and t_R that maximises the reduction in node impurity $i(t) - i(t_L) - i(t_R)$ is selected. Partitioning continues until the y values are constant in a node, or the node sample size is below a pre-specified threshold. Then the tree is pruned with the help of an independent test sample or by cross-validation and the subtree with the lowest estimated mean squared error is selected [59].

Cross-validation was applied to assess the optimal model complexity and minimise the risk of overfitting [60]. V-fold cross-validation (CV) involves the determination of a number of random sub-samples (in this case 10), which are extracted from the learning sample. Trees of a certain size are calculated V = 10 times, where, successively, one of the sub-samples is omitted in the calculation and used as the test sample in the cross-validation. Accordingly, each sub-sample is used V − 1 = 9 times in the learning sample and once as a test sample [61]. Furthermore, to avoid undesired model complexity, in our work the minimum size of each node which was divided into child nodes was defined as 15 cases. This means that any split of a node containing less than 15 cases was not accepted.

Two approaches were used during the construction of trees: in the first one, the process of the selection of exogenous variables participating in further partitions was automatic, i.e., each time the programme chose the variable leading to minimisation of the variance of the individual nodes and leaves of the tree and, thereby, to the improvement of "purity". In the second variant, interactive tree tools were used. The process of a tree construction is very similar to that of automatic regression trees, but, in this case, the researcher can select predictors participating in the partitions by himself/herself. For each predictor, the so-called "improvement measures" are calculated, which reflect the reduction of RSS (residual sum of squares), which can be achieved by using partitioning by each predictor. An improvement measure, as the basis for selecting the best predictor, is calculated in the same way as during the "automatic" construction of a tree (the aim is to minimise impurity), however, in this case the researcher can choose a predictor other than the one that most contributes to the increase of "purity". Selection of a different predictor (measure is not the highest) leads to an increase in "impurity", but is sometimes justified. This is the case, for instance, when a few variables have a similar value of improvement measure, but only one whose improvement of them is relatively easy to obtain (e.g., generally available in public statistics) and can have a broader application in practice.

A more detailed description and discussion of the theoretical concepts related to this data mining technique can be found in the reference sources [59,62]. In our project, all calculations were performed with the use of STATISTICA 13 software (StatSoft, Tulsa, OK, USA).

4. Results

In the first model describing factors differentiating the share of agricultural land in the surface area of the analysed municipalities of metropolitan areas, a tool for automated selection of external variables was used, resulting in a tree with minimum costs of wrong classifications. The program STATISTICA 13.0 used in the analysis generated a sequence of 33 regression trees with a varying degree of complexity, from a tree where all the elements belonged to one node, to a tree with very complex branches (Figure 2). Although tree no. 1 produces the smallest errors of object classification measured by cost of resubstitution, it shows relatively high costs of cross-validation, which indicates "overlearning of the tree" and lower predictive capabilities.

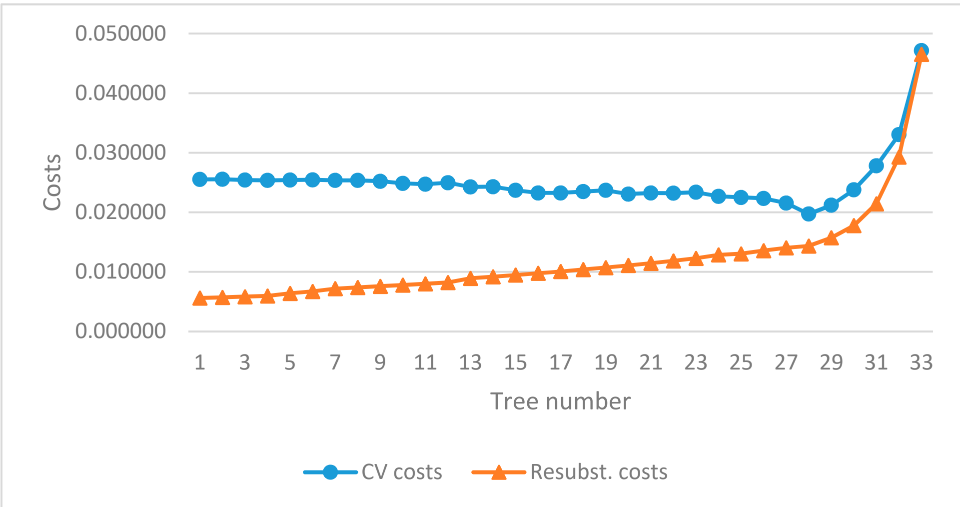

Figure 2. Chart of costs for a sequence of CART regression trees. Source: own elaboration.

Tree no. 28 was selected for further analyses. Its cross-validation costs are the lowest in the sequence of trees (Figure 3), and its moderate degree of complexity allowed it to preserve the desired interpretative and predictive capability.

In the Polish metropolitan areas selected for analyses, 49.9% of the overall surface area is occupied by agricultural land of individual farms, but a very large differentiation is visible here. In the light of automated selection of exogenous variables, the share of agricultural land was most impacted by entrepreneurship indicator. In municipalities where the number of non-agricultural sector enterprises is higher than 1960 entities/10,000 people of working age (threshold values for the individual external variables, in this case 1960 entities/10,000 people, are proposed by the model. They enable minimisation of the variance of child nodes), the share of agricultural land in the overall surface area is 29.9%, on average (node ID = 3). In the second group of municipalities, i.e., those with a lower entrepreneurship indicator (node ID = 2), agricultural land accounts for over 53.7% of the surface area.

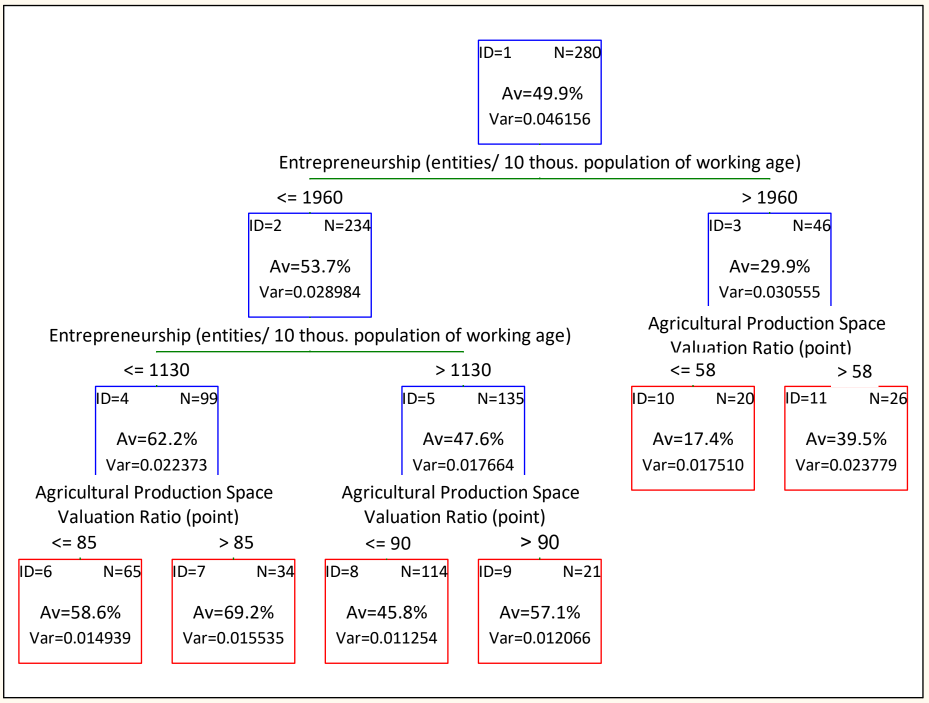

Figure 3. Model I—CART tree diagram, showing predictors influencing the share of agricultural land in the surface area of municipalities of metropolitan areas. Source: own elaboration.

Evaluating further partitions of the tree on the left-hand side, it is important to note that entrepreneurship is, again, the variable that best explains the differentiation in the percentage of agricultural land. It turns out that in municipalities where less than 1130 entities/10,000 people of working age carry out economic activity, agricultural land accounts for as much as 62.2%. If the indicator falls within the range of 1130–1960, then the average percentage of agricultural land is 47.6%.

The second very important variable determining the share of agricultural land in the surface area of the analysed municipalities is the Agricultural Production Space Valuation Ratio. This variable participated in partitions of nodes ID = 3, ID = 4, and ID = 5, and each time its relatively high value was accompanied by a relatively high share of agricultural land. Hence, it is important to note that good environmental conditions of agricultural production are conducive to the preservation of agricultural production, and thereby the percentage of agricultural land. Even in municipalities with a very large number of non-agricultural sector enterprises/10,000 people (node ID = 3), if the environmental conditions of agricultural production score above 58 points, almost 40% of the land (node ID = 11) is used for agricultural purposes whereas, otherwise, when the conditions are worse, the share is only 17.4%. Additionally, in municipalities with a low entrepreneurship indicator (node ID = 4), a lower Agricultural Production Space Valuation Ratio means a lower share of agricultural land, i.e., 58.6%, on average, whereas a high value of the Agricultural Production Space Valuation Ratio is accompanied by a very high share of agricultural land, i.e., almost 70%.

The very fact that two variables dominate in the model means that they are of very high importance in explaining the phenomenon being described and that they ensure the highest "purity" of the model at each stage of the tree partition. This does not mean, however, that the other variables are not important in the description of the phenomenon being analysed. One output of the CART procedure considers the importance of the predictors, which are ranked in descending order of their contribution to tree construction. To calculate a predictor importance score, CART calculates the improvement measure attributable to each variable in its role as a surrogate to the primary split (Table 5). The values of these improvements are summed over each node of the tree and scaled relative to the best performing variable. The variable with the highest sum of improvements is scored 100, and all other variables have lower scores ranging downwards towards zero [63].

Table 5. Importance of the predictors for the share of agricultural land in the surface area of a municipality.

Predictor	Predictor Category	Importance	Importance Category
Entrepreneurship	Socio-economic	100	Very important
Agricultural Production Space Valuation Ratio—APSVR	Environmental and fixed geographic feature	99	Very important
Large enterprises	Socio-economic	76	Important
Population density		71	Important
Net migration rate		67	Important
Type of municipality		61	Important
Commuting to work		50	Moderately important
Distance to the MA core	Environmental and fixed geographic feature	49	Moderately important
Protected natural areas		48	Moderately important
Surface area of municipality		45	Moderately important
Share of small farms	Structure of agriculture	38	Less important
Public roads	Environmental and fixed geographic feature	31	Less important
Unemployment rate	Socio-economic	26	Less important
Travel time to the MA core	Environmental and fixed geographic feature	23	Less important
Direct border with the MA core		21	Less important
Spatial planning		15	Less important
Share of economically strong farms	Structure of agriculture	13	Less important

Source: own elaboration.

The research has shown that the definitely most important predictors explaining the differentiation in the percentage of agricultural land were: entrepreneurship, environmental conditions of agricultural production, large enterprises, as well as predictors illustrating population density, net migration, and the type of municipality. The other variables, including the location of a municipality relative to the metropolis core, unemployment rate, and the structure of agriculture, turned out to be less important.

Based on the results of the model presented in Figure 3 and the ranking of predictor importance in Table 5, for the groups of municipalities concentrated in the different terminal nodes of Model I, the characteristics of selected variables used in the process of modelling are shown (Table 6).

Table 6. Characteristics of selected external variables for selected groups of municipalities.

Specification	>1960 Entities/10,000 Population of Working Age		<1130; 1960 => Entities/10,000 Population of Working Age		<=1130 Entities/10,000 Population of Working Age	
	APSVR <= 58 Points	APSVR > 58 Points	APSVR <= 90 Points	APSVR > 90 Points	APSVR <= 85 Points	APSVR > 85 Points
	Node ID = 10	Node ID = 11	Node ID = 8	Node ID = 9	Node ID = 6	Node ID = 7
Number of municipalities	20	26	114	21	65	34
Share of agricultural area (%)	17.4	39.5	45.8	57.1	58.6	69.2
Entrepreneurship (entities/10,000 population at the working age)	2631	2275	1484	1438	952	940
APSVR (points)	48	70	66	96	63	94
Population density (people/km^2)	983	1126	300	475	78	90
Net migration rate (person)	10.2	12.2	6.3	5.8	1.6	2.5
Commuting to work (person)	0.82	1.95	0.68	0.75	0.32	0.29

Table 6. *Cont.*

Specification	>1960 Entities/10,000 Population of Working Age		<1130; 1960 => Entities/10,000 Population of Working Age		<=1130 Entities/10,000 Population of Working Age	
	APSVR <= 58 Points	APSVR > 58 Points	APSVR <= 90 Points	APSVR > 90 Points	APSVR <= 85 Points	APSVR > 85 Points
	Node ID = 10	Node ID = 11	Node ID = 8	Node ID = 9	Node ID = 6	Node ID = 7
Type of municipality						
Urban	17	12	12	4	0	0
Urban-rural	0	7	47	7	6	4
Rural	3	7	55	10	59	30
Distance to the MA core (km)	29.2	19.2	39.9	27.4	44.4	30.5
Direct border with the MA						
Core	0	5	0	1	0	0
Yes	12	15	29	11	2	7
No	8	6	85	9	63	27
Overall average surface area of municipality (ha)	4564	11,304	11,584	9769	12,110	9554

Source: own elaboration.

In accordance with the results of Model I (Figure 3), the percentage of agricultural land in municipalities of metropolitan areas is mainly determined by entrepreneurship and the quality of agricultural production space. However, as regards the principle saying that the higher the entrepreneurship indicator, the smaller share of agricultural land, the research has shown that agriculture is more likely to maintain the area it occupies if very good environmental conditions of agricultural production are present. It is important to note that in the municipalities concentrated in nodes ID = 6 and ID = 7, as well as in nodes ID = 8 and ID = 9 (Table 6), the entrepreneurship indicator is very similar, and, in both cases, in municipalities with better environmental conditions of agricultural production the percentage of agricultural land is higher by over 10 percentage points than in municipalities with worse conditions of agricultural production (node ID = 7 versus node ID = 6 and node ID = 9 versus node ID = 8). The significant importance of environmental conditions of agricultural production is particularly visible in the group of municipalities with a medium entrepreneurship indicator and very good conditions of agricultural production (node ID = 9). Although there is a very high pressure of urbanisation manifested in a high population density (475 people/km^2) and a relatively high net migration (5.8 people/1000 people), and most municipalities are located in the immediate proximity of the core, agricultural land occupies as much as 57.1% of the overall surface area. This is a value comparable with municipalities concentrated in node ID = 6 where there is a relatively low entrepreneurship (below 1130 entities/10,000 population at the working age), low urbanisation pressure, and as many as 59/65 municipalities are located in rural areas mainly on the outskirts of metropolitan areas. Hence, it appears that very good conditions of agricultural production (in our research above 90 points on the APSVR) can hinder the spatial development of cities.

Proceeding to analysis of municipalities concentrated in nodes ID = 10 and ID = 11, i.e., municipalities with a very high entrepreneurship indicator, we should highlight that the share of agricultural land differs greatly: 35.5% vs. 17.4%. The model shows that environmental conditions of agricultural production are the main factor differentiating these municipalities, which is consistent with the adopted line of reasoning (theory of economic rents) and the academic literature. However, other factors may play a significant role here. It should be pointed out that, among the municipalities concentrated in node ID = 11, i.e., municipalities with relatively good environmental conditions, there are five cores of metropolitan areas that are more than five times larger than the other municipalities, and these cities very dynamically increased their surface area in the past by absorbing the surrounding villages [64]. Thus, the relatively large percentage of agricultural land in the cores of metropolitan areas may be a result of their large surface area, as they have large reserves of land that the non-agricultural sector is unable to use. The opposite can be observed in the case of small urban municipalities (17/20 municipalities) concentrated in node ID = 10, which are characterised by a relatively small percentage of agricultural land. Apart from differences in the surface area of municipalities concentrated in nodes ID = 10 and ID = 11, another important variable determining the share of agricultural land can be the status/type of municipalities. In urban municipalities located in the outer zone of metropolitan areas (towns), processes of urbanisation are most intensive, there is compact development, advanced technical infrastructure, and a small percentage of people employed in agriculture. As a result, towns, especially smaller ones, i.e., with up to 20,000 people, have a smaller percentage of agricultural land in the overall surface area [15]. In accordance with agricultural production location theories, higher population density and positive net migration should be associated with a smaller share of agricultural land, but in the case of the municipalities concentrated in nodes ID = 6 and ID = 7, there is the opposite relationship. Thus, in the first model, constructed strictly on the basis on statistical criteria, not all relationships may have been captured.

Since Model I (Figure 3) is very simple, and substantive considerations (as shown above) indicate a range of alternative variables which, if used, may cast a different light on the mechanisms of competing for land, the decision was taken to build an alternative regression tree using the technique of interactive regression trees (Figure 4). Although the entrepreneurship variable ensures the highest "purity" of the model (as shown in Model I), it was decided that the first split of the tree into nodes ID = 2 and ID = 3 will be performed using the variable "type of municipality", which, at this stage of construction of a tree, occupied third position in the improvement measure ranking (at each stage of construction of a tree the model estimates improvement measures). This variable is substantively justified and indicates logical partitions, as in urban municipalities (small cities) the share of agricultural land was 32.6% on average, and in the remaining ones, i.e., rural and urban-rural municipalities, it accounted for around 53.1%. Proceeding to the analysis of urban municipalities, including the cores of metropolitan areas (node ID = 2), it is important to note that in municipalities where net migration was positive, the share of agricultural land is 24.5%, on average, while in municipalities with a negative net migration it is 44.4%. Thus, it turns out that in urban municipalities which, for various reasons, have not attracted new inhabitants and, thus, have not recorded an increase in the demand for land to be used for non-agricultural purposes, the share of agricultural land remains quite high. This situation referred mainly to urban municipalities (cities) located relatively far from the cores of metropolises, but it was also visible in the cores with a large surface area: Lublin, Poznan, and the Tricity, where processes of suburbanisation occurred during the period analysed.

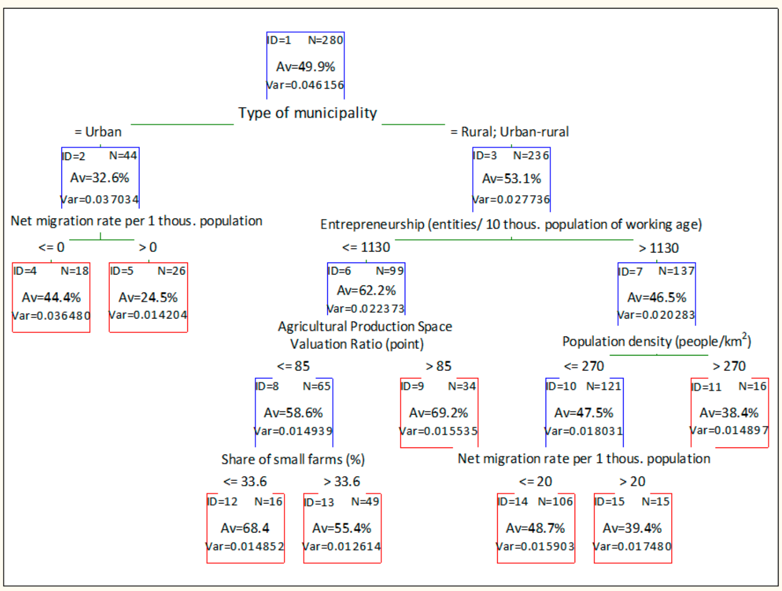

Figure 4. Model II—CART tree diagram, showing predictors influencing the share of agricultural land in the surface area of municipalities. Source: own elaboration.

Analysing the share of agricultural land in rural and urban-rural municipalities (node ID = 3), it can be noted that the entrepreneurship indicator is definitely the most important factor. Its relatively high values, i.e., above 1130 entities/10,000 people, are associated with a share of agricultural land of 46.5%, whereas its relatively low values are accompanied by a share of agricultural land of over 62.2%. Further partitions of node ID = 7, i.e., in the group of rural and urban-rural municipalities characterised by a relatively high level of entrepreneurship, show that the share of agricultural land depends on the population density (node ID = 10 and ID = 11) and net migration (node ID = 14 and ID = 15). With population density greater than 270 people/km^2, the share of agricultural land in the analysed municipalities is 38.4%, and if the population density is below 270 people/km^2, it is around 47.5%. However, also in municipalities having a lower population density (node ID = 10) with a relatively high net migration, the share of agricultural land is 39.4% (node ID = 15). Thus, it is important to note that a relatively high pressure of urbanisation, manifested in a high population density or a positive net migration, is accompanied by a low share of agricultural land.

In rural and urban-rural municipalities, where there is a relatively low number of non-agricultural sector enterprises, similarly to Model I, the share of agricultural land in the surface area of a municipality is determined by the quality of agricultural production space. Very good environmental conditions (node ID = 9) come with a high share of agricultural land, i.e., over 69% of the overall surface area of municipalities. In the case of worse environmental conditions for agricultural production, especially in municipalities where the share of small farms is above 34% (ID = 13), the share of agricultural land in the surface area of a municipality is only 55.4%. Thus, it can be noted that in the case of municipalities with a low level of entrepreneurship accompanied by a relatively low population density and low net migration, similarly as in Model I, the share of agricultural land is mainly determined by the quality of agricultural production space.

Model II was based on the same external variables as Model I, therefore, the ranking of predictor importance is very similar for both the models and, for that reason, will not be presented. The alternative model was based on expert knowledge, so the variables maximising the model's "purity" were not always the ones that were selected. When interpreting resubstitution costs of the selected regression tree model, we should compare them to the variance characterising the exogenous variable, which, in the first node, was 0.046156. Resubstitution costs represent a kind of measure of the reduction of "impurity", and the greater the decrease in these costs, the better the model. In our case, the partitions used in Model II enabled reduction of the variance of the average share of agricultural land in the surface area of a municipality by around 64% (to the level of 0.016415). In the case of Model I, automatic partitions enabled the reduction of the variance by around 69% (to the level of 0.014344). Reduction of the overall variance to 0 would mean that a model perfectly represents the reality, therefore, the improvement achieved in the quality of the match of the models to empirical data at a level of almost 70% can be regarded as satisfactory. The models presented explain the differentiation in the share of agricultural land in the overall surface area of municipalities of metropolitan areas quite well.

5. Discussion

The presented models address factors determining the differentiation in the share of agricultural land in the overall surface area of the analysed municipalities of metropolitan areas. The analyses showed that the most important factors determining the share of agricultural land in the surface area of the analysed municipalities were definitely the entrepreneurship indicator and the agricultural production space valuation ratio. These variables dominated partitions of regression tree no. 1 (Figure 3) and occupied first and second positions, respectively, in the ranking of predictors (Table 5). The analyses confirmed that the differentiation in the share of agricultural land in metropolitan areas is, to a large extent, determined by economic rents. The first of the variables, i.e., the entrepreneurship indicator, which is measured by calculating the number of non-agricultural sector enterprises per 10,000 people, represents "urban land use", i.e., economic rents generated by the non-agricultural sector. The second variable directly refers to economic rent formulated by Ricardo and reflects rents for agricultural land use depending on land fertility. In accordance with models developed by Alonso [26] and Konagaya [35], higher economic rents generated by the non-agricultural sector (urban land use) lead to the shrinking of land used for agricultural purposes, which is pushed out of areas/municipalities where the non-agricultural sector is developing. In our research, we noticed that the percentage of agricultural land decreases as the number of enterprises increases. Economic and employment growth lead to urban land development [8], therefore, in municipalities with a large number of enterprises the percentage of agricultural land is relatively low. Similar results are presented by Deng et al. [10], who analyse the impact of the increase of industrial GDP on the scale of farmland conversion. They note that an increase in GDP from industry leads to a decrease in the surface area of agricultural land. They also argue that the development of industrial sector enterprises is connected with expansion of warehousing space, infrastructure, etc., which leads to a decrease in the surface area of land used for agricultural purposes. Enterprises need space in order to grow, but of equal importance is the fact that they generate jobs, and attractive job offers attract new inhabitants (more land is needed for house-building) and make the owners of small farms abandon agricultural production to seek employment in the city. As a result, especially in municipalities with a high degree of agricultural fragmentation (Model II; ID = 13), the share of agricultural land in the overall surface area of cities is relatively low. Sroka [15] notes that such municipalities see very dynamic processes of shifts away from agriculture; farmland is set aside in the initial phase to be later converted to non-agricultural purposes, i.e., house-building and business development. A negative impact of the non-agricultural labour market on agricultural activity is noted, among others, by Wästfelt and Zhang [30] and Pölling et al. [65], who point out that it is especially small farm owners and part-time farmers who give up agricultural activity.

The large significance of the "entrepreneurship" variable is further increased by the impact of other variables, including a relatively large population density, positive net migration, and commuting to work. Thus, it turns out that the development of non-agricultural enterprises is accompanied by urbanisation pressure. The importance of these variables was presented in Table 5 and Model II, which indicated that a relatively high net migration and high population density both come with a relatively low share of agricultural land (Figure 4). An increase in the number of inhabitants requires that land is allocated for house building, as well as technical and social infrastructure. As Alonso [26] has already noted, higher economic rents generated by the development of house building contribute to a decrease in the area of land used for agricultural purposes. Also Mazzocchi et al. [9] noted when researching the metropolitan area of Milan that population density measures the pressure of the population on an area, in itself a negative factor for agriculture as it represents a threat to free space given the probable positive relation between population density and the demand for new homes and services. Huang et al. [8] and Wu and Zhang [66] arrived at similar conclusions. Meanwhile, Zasada et al. [37] stress that an increase in population density and migration contribute to production extensification and land set-aside. In the conditions of Eastern and Central European countries, the setting aside of land usually leads to a total abandonment of production, with farmland becoming subject to speculation [41].

The academic literature very often stresses that environmental conditions of agricultural production determine the productivity and profitability of agriculture [19,67,68]. Our research has shown that, in municipalities with better environmental conditions, the share of agricultural land is relatively high. This is because higher productivity of agricultural land, especially when the land is intensively cultivated, results in higher economic rent from every unit of surface area. In metropolitan areas, which often occupy land of very good quality, horticulture and vegetable production develops, which, even in the case of small areas, ensures relatively high income and the preservation of farming [69]. In such areas, land users less often give up agricultural production and convert land to non-agricultural uses. This relationship is also shown in other authors' research. For instance, Baumann et al. [70] note, while dealing with agriculture in Ukraine, that soil type and topography explain the spatial heterogeneity of farmland abandonment. They stress that when yield potential is high (and economic rent is, too), the likelihood of farmland abandonment falls. Similar conclusions are drawn by Xie, at al. [40], as well as Gellrich and Zimmermann [39]. Although the research cited in the paper did not analyse the process of farmland conversion, it should be stressed that set-aside usually precedes permanent abandonment of farmland. Under the conditions found in Poland, the relatively high share of agricultural land in municipalities with good environmental conditions is also an effect of more stringent restrictions imposed by legal regulations regarding farmland protection. Farmland of the best quality can only be converted to non-agricultural uses with the consent of the Minister of Agriculture. Moreover, this requires the payment of quite high fees (which may even be in excess of 100,000 euros per 1 ha) and involves long procedures [13]. As a result, near such cities as Krakow, Lublin, or Wroclaw, where soil conditions are among the best in the country, there are still municipalities with an agricultural profile. The local people are not interested in selling their land, and high specialisation of production enables the generation of comparable incomes.

Contrary to other authors [8–10], we have noted that predictors illustrating the distance, travel time, or the location relative to the core (direct vicinity) turned out to be of relatively little importance. However, it should be stressed that the studies cited in the paper analysed the impact of location on the shrinkage of agricultural land, whereas our research deals, among other things, with the impact of location on the percentage of agricultural land. Thus, our findings do not contradict the findings presented by other authors, but only show that proximity alone is not translated into a small share of agricultural land. In the analysed metropolitan areas, in as many as 29 out of 75 municipalities that border on the cores, and are thereby located within a small distance from the core, the percentage of agricultural land is over 50% of the overall surface area, whereas in 10 municipalities it is over 70%. This refers to municipalities with very good environmental conditions. It can, thus, be concluded that good environmental conditions hinder processes of urban expansion.

6. Conclusions

The analyses were conducted to assess the factors having an impact on the share of agricultural land in municipalities of selected Polish metropolitan areas. The research was based on the concept of economic rents, which was used to explain the mechanism of competing for land. The models of regression trees used in the paper are mainly descriptive in character, i.e., they aim to describe and present mechanisms by which socio-economic factors, environmental and fixed geographic features, and the structure of agriculture in the examined population have an impact on the share of agricultural land in the surface area of the analysed municipalities.

The research has shown that, in the Polish metropolitan areas selected for analysis, 49.9% of land is, on average, occupied by agricultural land of individual farms. The factors that turned out to be the strongest determinants of the share of agricultural land in the surface area of the analysed municipalities were the entrepreneurship indicator, which is measured by the number of enterprises per 10,000 people of working age, and the Agricultural Production Space Valuation Ratio, which indicates the quality of soil, climate for farming, water conditions, and topography. It was found that with an increase in the entrepreneurship indicator, the share of agricultural land falls, and the high quality of agricultural production space comes with a relatively high share of agricultural land in the surface area of the analysed municipalities. It has also been noted that the spatial importance of agricultural land is determined by population density indices, the migration index, and commuting to work. Higher values of these indices translate into a lower share of agricultural land. To live and work, humans need space, so population growth, high population density, migration, as well as economic and employment growth, lead to a decrease in the share of agricultural land.

In metropolitan areas, agriculture competes for land with urban land uses, which generate higher economic rents. That is why in municipalities with high urbanisation pressure the share of agricultural land was relatively low. However, the research has shown that even in highly-urbanised municipalities, including those located at the border with the cores of metropolitan areas, it is possible to maintain a large percentage of agricultural land. However, this requires an appropriately high quality of environmental conditions of agricultural production. In such municipalities, economic rent arising from the high fertility of soil, and also the high fees associated with taking plots of land out of agricultural production (up to 100.000 euro/ha) can exceed urban rents and, thus, the proportion of farmland to land used for non-agricultural purposes is in favour of farmland. However, in these municipalities, too, these relations will be subject to correction, and the less profitable agriculture becomes, the larger the percentage of land allocated for non-agricultural purposes will be. The identified mechanisms of competition for land can be useful in the development of land use policy in certain post-communist countries (especially in Europe) and countries showing dynamic economic growth, as well as countries which are undergoing, or have undergone, transformation of the land management model. We have shown that in Poland, i.e., a country with a relatively poor protection of the agricultural land market, the way that land is used in metropolitan areas is determined by the amount of economic rents. The mechanism of economic rents prefers uses that bring high economic rents, therefore, maintenance of an appropriate share of agricultural land, especially in areas undergoing processes of urbanisation, will only be possible where the state plays an active role and uses instruments for farmland protection.

The availability of the data ensures the replicability of the analysis and reproduction of the results, while also providing maps that are easy to read. The model presented may also be transformed into a prognostic model, and the results of the analysis may prove useful in identifying municipalities in metropolitan areas, which may be affected by the process of the conversion of land used for agricultural purposes in the future. In particular, the very simple Model I makes it possible to estimate how, under existing (unchanging) natural conditions of agricultural production, the share of land used for agricultural purposes will evolve as the number of enterprises in the non-agricultural sector changes. Such knowledge allows one to plan additional courses of action to be taken to protect agricultural areas, for example, within the framework of adopted spatial development plans. Of course, it will be very difficult to stop processes of urbanisation entirely, but at least it may be possible to protect a part of agricultural land that is particularly suitable for agricultural production and areas of land, which fulfil other functions, such as areas of natural beauty, which form ecological corridors or clean-air corridors that allow cities to "breathe".

When analysing these findings, it should be borne in mind that the presented models are 70% better at explaining the differentiation in the share of agricultural land than a zero model, i.e., a model without external variables. Due to the lack of reliable materials, the research has not taken into account farmland held in other forms of land ownership than individual farmlands. This may somewhat distort the picture of land use, particularly in cities where a relatively large part of land is owned by enterprises and municipal authorities.

Supplementary Materials: The following are available online at http://www.mdpi.com/2071-1050/10/3/837/s1.

Acknowledgments: We thank the editor and the reviewers for their valuable comments and detailed suggestions to improve the paper. Further, we also acknowledge the support by the National Science Centre, Poland under project no 2016/21/D/HS4/00264.

Author Contributions: Wojciech Sroka developed the original idea and contributed to the research design. Jarosław Mikołajczyk and Tomasz Wojewodzic were responsible for data collection. Bogusława Kwoczyńska contributed to the research design and provided guidance. Wojciech Sroka wrote the paper. All authors have read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kizos, T.; Tsilimigkas, G.; Karampela, S. What Drives Built-Up Area Expansion on Islands? Using Soil Sealing Indicators to Estimate Built-Up Area Patterns on Aegean Islands, Greece. *Tijdschr. Econ. Soc. Geogr.* **2017**, *108*, 703–886. [CrossRef]

2. Piorr, A.; Ravetz, J.; Tosics, I. *Peri-Urbanisation in Europe: Towards European Policies to Sustain Urban-Rural Futures*; Forest & Landscape, University of Copenhagen: Copenhagen, Denmark, 2011; pp. 30–42, ISBN 8779035345.
3. Krätke, S. Metropolisation of the European economic territory as a consequence of increasing specialisation of urban agglomerations in the knowledge economy. *Eur. Plan. Stud.* **2007**, *15*, 1–27. [CrossRef]
4. Smętkowski, M.; Jałowiecki, B.; Gorzelak, G. *Obszary Metropolitalne w Polsce. Problemy Rozwojowe i Delimitacja*; Raporty i Analizy EUROREG 1; Centrum Europejskich Studiów Regionalnych i Lokalnych EUROREG: Warszawa, Poland, 2009; pp. 8–26.
5. Zasada, I. Multifunctional peri-urban agriculture—A review of societal demands and the provision of goods and services by farming. *Land Use Policy* **2011**, *28*, 639–648. [CrossRef]
6. Lovell, S.T. Multifunctional Urban Agriculture for Sustainable Land Use Planning in the United States. *Sustainability* **2010**, *2*, 2499–2522. [CrossRef]
7. Busko, M.; Szafranska, B. Analysis of Changes in Land Use Patterns Pursuant to the Conversion of Agricultural Land to Non-Agricultural Use in the Context of the Sustainable Development of the Malopolska Region. *Sustainability* **2018**, *10*, 136. [CrossRef]
8. Huang, D.; Jin, H.; Zhao, X.; Liu, S. Factors influencing the conversion of arable land to urban use and policy implications in Beijing, China. *Sustainability* **2015**, *7*, 180–194. [CrossRef]
9. Mazzocchi, C.; Sali, G.; Corsi, S. Land use conversion in metropolitan areas and the permanence of agriculture: Sensitivity Index of Agricultural Land (SIAL), a tool for territorial analysis. *Land Use Policy* **2013**, *35*, 155–162. [CrossRef]
10. Deng, X.; Huang, J.; Rozelle, S.; Zhang, J.; Li, Z. Impact of urbanization on cultivated land changes in China. *Land Use Policy* **2015**, *45*, 1–7. [CrossRef]
11. Su, S.L.; Jiang, Z.L.; Zhang, Q.; Zhang, Y.A. Transformation of agricultural landscapes under rapid urbanization: A threat to sustainability in Hang-Jia-Hu region, China. *Appl. Geogr.* **2011**, *31*, 439–449. [CrossRef]
12. Lorens, P.; Martyniuk-Pęczek, J. *Zarządzanie Rozwojem Przestrzennym Miast*; Wydawnictwo Urbanista: Gdańsk, Poland, 2010; pp. 16–42, ISBN 978-83-89649-30-0.
13. Krzyk, P.; Tokarczuk, T.; Heczko-Hyłowa, E.; Ziobrowski, Z. *Obszary Rolne Jako Element Struktury Przestrzennej Miast—Problemy Planistyczne*; Instytut Rozwoju Miast: Kraków, Poland, 2013; pp. 48–69. ISBN 978-83-89440-72-3.
14. Poławski, Z.F. Zmiany użytkowania ziemi w Polsce w ostatnich dwóch stuleciach. *Teledetekcja Środowiska* **2009**, *42*, 74–75.
15. Sroka, W. Resources and use of agricultural land in Polish cities according to chosen theories of location of agricultural production. *Acta Sci. Pol. Oecon.* **2015**, *14*, 135–147.
16. Hagedorn, K. Property rights reform on agricultural land in Central and Eastern Europe. *Q. J. Int. Agric.* **2004**, *43*, 409–438.
17. Czyżewski, A.; Majchrzak, A. Regulations for the Polish agricultural land market compared with other EU member states: Implications for market efficiency. *Stud. Ekon.* **2014**, *2*, 175–195.
18. Qian, J.; Peng, Y.; Luo, C.; Wu, C.; Du, Q. Urban land expansion and sustainable land use policy in Shenzhen: A case study of China's rapid urbanization. *Sustainability* **2016**, *8*, 16. [CrossRef]
19. Diogo, V.; Koomen, E.; Kuhlman, T. An economic theory-based explanatory model of agricultural land-use patterns: The Netherlands as a case study. *Agric. Syst.* **2015**, *139*, 1–16. [CrossRef]
20. Van der Veen, A.; Otter, H.S. Land use changes in regional economic theory. *Environ. Model. Assess.* **2001**, *6*, 145–150. [CrossRef]
21. Nalej, M. Agricultural land cover changes in metropolitan areas of Poland for the period 1990–2012. *Misc. Geogr.* **2016**, *20*, 39–45. [CrossRef]
22. Abrantes, P.; Fontes, I.; Gomes, E.; Rocha, J. Compliance of land cover changes with municipal land use planning: Evidence from the Lisbon metropolitan region (1990–2007). *Land Use Policy* **2016**, *51*, 120–134. [CrossRef]
23. Van Rij, E.; Dekkers, J.; Koomen, E. Analysing the success of open space preservation in the Netherlands: The midden-delfland case. *Tijdschr. Econ. Soc. Geogr.* **2008**, *99*, 115–124. [CrossRef]
24. Starrett, D. Market allocations of location choice in a model with free mobility. *J. Econ. Theory* **1978**, *17*, 21–37. [CrossRef]

25. Fujita, M. The evolution of spatial economics: From Thünen to the new economic geography. *Jpn. Econ. Rev.* **2010**, *61*, 1–32. [CrossRef]
26. Alonso, W. *Location and Land Use*; Harvard University Press: Cambridge, MA, USA, 1964; ISBN 9780674730854.
27. Lambin, E.F.; Rounsevell, M.D.A.; Geist, H.J. Are agricultural land-use models able to predict changes in land-use intensity? *Agric. Ecosyst. Environ.* **2000**, *82*, 321–331. [CrossRef]
28. Ustaoglu, E.; Castillo, C.P.; Jacobs-Crisioni, C.; Lavalle, C. Economic evaluation of agricultural land to assess land use changes. *Land Use Policy* **2016**, *56*, 125–146. [CrossRef]
29. Czyżewski, B.; Matuszczak, A. A new land rent theory for sustainable agriculture. *Land Use Policy* **2016**, *55*, 222–229. [CrossRef]
30. Wästfelt, A.; Zhang, Q. Reclaiming localisation for revitalising agriculture: A case study of peri-urban agricultural change in Gothenburg, Sweden. *J. Rural Stud.* **2016**, *47*, 172–185. [CrossRef]
31. O'Kelly, M.; Bryan, D. Agricultural location theory: Von Thünen's contribution to economic geography. *Prog. Hum. Geogr.* **1996**, *20*, 457–475. [CrossRef]
32. Kellerman, A. Determinants of rent from agricultural land around metropolitan areas. *Geogr. Anal.* **1978**, *10*, 1–12. [CrossRef]
33. Sinclair, R. Von Thünen and urban sprawl. *Ann. Assoc. Am. Geogr.* **1976**, *57*, 72–87. [CrossRef]
34. Wegener, M. Overview of Land Use Transport Models. In *Handbook of Transport Geography and Spatial Systems*; Hensher, D.A., Button, K.J., Haynes, K.E., Stopher, P.R., Eds.; Handbooks in Transport; Elsevier: Amsterdam, The Netherlands, 2004; Volume 5, pp. 127–146.
35. Konagaya, K. The Generalised Thunen Alonso Model for Land use Change in Sumatra Island. *Geogr. Environ. Model.* **1999**, *3*, 145–162.
36. Mitsuda, Y.; Ito, S. A review of spatial-explicit factors determining spatial distribution of land use/land-use change. *Landsc. Ecol. Eng.* **2011**, *7*, 117–125. [CrossRef]
37. Zasada, I.; Fertner, C.; Piorr, A.; Nielsen, T.S. Peri-urbanisation and multifunctional adaptation of agriculture around Copenhagen. *Dan. J. Geogr.* **2011**, *111*, 59–72. [CrossRef]
38. Ravetz, J.; Fertner, C.; Nielsen, T.S. The dynamics of peri-urbanization. In *Peri-Urban Futures: Scenarios and Models for Land Use Change in Europe*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 13–44, ISBN 978-3-642-30528-3.
39. Gellrich, M.; Zimmermann, N.E. Investigating the regional-scale pattern of agricultural land abandonment in the Swiss mountains: A spatial statistical modelling approach. *Landsc. Urban Plan.* **2007**, *79*, 65–76. [CrossRef]
40. Xie, H.; Wang, P.; Yao, G. Exploring the dynamic mechanisms of farmland abandonment based on a spatially explicit economic model for environmental sustainability: A case study in Jiangxi Province, China. *Sustainability* **2014**, *6*, 1260–1282. [CrossRef]
41. Grădinaru, S.R.; Iojă, C.I.; Onose, D.A.; Gavrilidis, A.A.; Pătru-Stupariu, I.; Kienast, F.; Hersperger, A.M. Land abandonment as a precursor of built-up development at the sprawling periphery of former socialist cities. *Ecol. Indic.* **2015**, *57*, 305–313. [CrossRef]
42. Pölling, B.; Mergenthaler, M.; Lorleberg, W. Professional urban agriculture and its characteristic business models in Metropolis Ruhr, Germany. *Land Use Policy* **2016**, *58*, 366–379. [CrossRef]
43. Kopsidis, M.; Wolf, N. Agricultural productivity across Prussia during the industrial revolution: A Thünen perspective. *J. Econ. Hist.* **2012**, *72*, 634–670. [CrossRef]
44. Samorząd Województwa Lubelskiego. *Plan Zagospodarowania Przestrzennego Województwa Lubelskiego*; Urząd Marszałkowski Województwa Lubelskiego: Lublin, Poland, 2015; Available online: https://umwl.bip.lubelskie.pl/upload/pliki//2Zal.1_tekst_PZPWL.pdf (accessed on 2 September 2017).
45. Wielkopolskie Biuro Planowania Przestrzennego. *Studium Ochrony i Kształtowania Krajobrazu W Poznańskim Obszarze Metropolitalnym*; Wielkopolskie Biuro Planowania Przestrzennego: Poznań, Poland, 2012; Available online: http://www.wbpp.poznan.pl/opracowania/krajobraz/krajobraz.html (accessed on 2 September 2017).
46. Urząd Statystyczny w Krakowie. *Krakowski Obszar Metropolitalny w Latach 2011–2015*; Urząd Statystyczny w Krakowie: Kraków, Poland, 2016. Available online: http://krakow.stat.gov.pl/publikacje-i-foldery/inne-opracowania/krakowski-obszar-metropolitalny-w-latach-2011-2015,7,2.html (accessed on 2 September 2017).

47. Pomorskie Biuro Planowania Regionalnego. *Plan Zagospodarowania Przestrzennego Obszaru Metropolitalnego Gdańsk-Gdynia-Sopot 2030*; Pomorskie Biuro Planowania Regionalnego: Gdańsk, Poland, 2016; Available online: http://www.metropoliagdansk.pl/upload/files/PZPOMG-G-S%202030(1).pdf (accessed on 2 September 2017).
48. Mazowieckie Biuro Planowania Regionalnego w Warszawie. *Studium Planu Zagospodarowania Przestrzennego Obszaru Metropolitalnego Warszawy*; Mazowieckie Biuro Planowania Regionalnego w Warszawie: Warszawa, Poland, 2010; Available online: http://user.siskom.waw.pl/userftp/materialy/mat_dostepne/planistyka/studium_pzp_om_warszawy/studium_pzpomw_2010.pdf (accessed on 2 September 2017).
49. Instytut Rozwoju Terytorialnego. *Plan Zagospodarowania Przestrzennego Województwa Dolnośląskiego. Perspektywa 2020*; Instytut Rozwoju Terytorialnego: Wrocław, Poland, 2014; Available online: http://www.pzpwd.pl/download/pzpwd.pdf (accessed on 2 September 2017).
50. Central Statistical Office of Poland (GUS). *Rolnictwo na Terenach Miejskich w Województwie Mazowieckim*; Urząd Statystyczny w Warszawie: Warszawa, Poland, 2013; pp. 29–30, ISBN 978-83-63246-21-1.
51. Witek, T. *Waloryzacja Rolniczej Przestrzeni Produkcyjnej Polski Według Gmin*; IUNG: Puławy, Poland, 1994; pp. 166–248.
52. Wiland, M. Spekulacja gruntami—Budownictwo mieszkaniowe—Depopulacja. In *Mieszkalnictwo w Roku Rodziny: Raport CBOS oraz Ekspertyzy i Opinie w Sprawie Sytuacji Mieszkaniowej Polaków*; Habitat for Humanity; Fundacja Bezdomnych: Warszawa, Poland, 2013; pp. 58–60.
53. Rouha-Mülleder, C.; Iben, C.; Wagner, E.; Laaha, G.; Troxler, J.; Waiblinger, S. Relative importance of factors influencing the prevalence of lameness in Austrian cubicle loose-housed dairy cows. *Prev. Vet. Med.* **2009**, *92*, 123–133. [CrossRef] [PubMed]
54. Rokach, L.; Maimon, O. *Data Mining with Decision Trees—Theory and Applications*; World Scientific Publishing: Hackensack, NJ, USA, 2008; pp. 81–83, ISBN 978-9814590075.
55. Kuhn, L.; Page, K.; Ward, J.; Worrall-Carter, L. The process and utility of classification and regression tree methodology in nursing research. *J. Adv. Nurs.* **2014**, *70*, 1276–1286. [CrossRef] [PubMed]
56. Breiman, L.; Friedman, J.H.; Olshen, R.A.; Stone, C.J. *Classification and Regression Trees*; Wadsworth & Brooks: Monterey, CA, USA, 1984; ISBN 0412048418.
57. Dacko, M.; Szajdecka, K. Multifaceted analysis of the process of price developments on the local real estate market by means of the regression trees method (CandRT). *Acta Sci. Pol. Oecon.* **2015**, *14*, 27–38.
58. Topal, M.; Aksakal, V.; Bayram, B.; Yağanoğlu, A.M. An analysis of the factors affecting birth weight and actual milk yield in Swedish red cattle using regression tree analysis. *J. Anim. Plant Sci.* **2010**, *20*, 63–69.
59. Loh, W.Y.; Zheng, W. Regression trees for longitudinal and multiresponse data. *Ann. Appl. Stat.* **2013**, *7*, 495–522. [CrossRef]
60. Ronowicz, J.; Thommes, M.; Kleinebudde, P.; Krysiński, J. A data mining approach to optimize pellets manufacturing process based on a decision tree algorithm. *Eur. J. Pharm. Sci.* **2015**, *73*, 44–48. [CrossRef] [PubMed]
61. Dacko, M.; Zając, T.; Synowiec, A.; Oleksy, A.; Klimek-Kopyra, A.; Kulig, B. New approach to determine biological and environmental factors influencing mass of a single pea (*Pisum sativum* L.) seed in Silesia region in Poland using a CART model. *Eur. J. Agron.* **2016**, *74*, 29–37. [CrossRef]
62. Dębska, B.; Guzowska-Świder, B. Decision trees in selection of featured determined food quality. *Anal. Chim. Acta* **2011**, *705*, 261–271. [CrossRef] [PubMed]
63. Ives, C.D.; Kendal, D. Values and attitudes of the urban public towards peri-urban agricultural land. *Land Use Policy* **2013**, *34*, 80–90. [CrossRef]
64. Szymańska, D.; Grzelak-Kostulska, E.; Hołowiecka, B. Zmiany powierzchni i gęstości zaludnienia miast Polski w latach 1960–2003. In *Kierunki Przekształceń Struktury Gospodarczej i Społeczno-Demograficznej Miast*; Słodczyk, J., Szafranek, E., Eds.; Wydawnictwo Uniwersytetu Opolskiego: Opole, Poland, 2006; pp. 341–353. ISBN 83-7395-189-X.
65. Pölling, B.; Sroka, W.; Mergenthaler, M. Success of urban farming's city-adjustments and business models—Findings from a survey among farmers in Ruhr Metropolis, Germany. *Land Use Policy* **2017**, *69*, 372–385. [CrossRef]
66. Wu, K.Y.; Zhang, H. Land use dynamics, built-up land expansion patterns, and driving forces analysis of the fast-growing Hangzhou metropolitan area, eastern China (1978–2008). *Appl. Geogr.* **2012**, *34*, 137–145. [CrossRef]

67. Ewert, F.; Rounsevell, M.D.A.; Reginster, I.; Metzger, M.J.; Leemans, R. Future scenarios of European agricultural land use: I. Estimating changes in crop productivity. *Agric. Ecosyst. Environ.* **2005**, *107*, 101–116. [CrossRef]
68. Müller, D.; Zeller, M. Land use dynamics in the central highlands of Vietnam: A spatial model combining village survey data with satellite imagery interpretation. *Agric. Econ.* **2002**, *27*, 333–354. [CrossRef]
69. Sroka, W.; Pölling, B. The Potential and Significance of urban Agriculture on the basis of the Ruhr Metropolis and the Upper Silesian Metropolis. *Probl. World Agric.* **2015**, *15*, 180–193.
70. Baumann, M.; Kuemmerle, T.; Elbakidze, M.; Ozdogan, M.; Radeloff, V.C.; Keuler, N.S.; Hostert, P. Patterns and drivers of post-socialist farmland abandonment in Western Ukraine. *Land Use Policy* **2011**, *28*, 552–562. [CrossRef]