FORLand Technical Paper 02 (2021)

Quantification of Ownership Concentration from Cadastral Records of Agricultural Land in Märkisch-Oderland

Daniel Müller, Philippe Rufin, Marcel Schwieder

Concentration of ownership of agricultural land in fewer hands has often been highlighted as a side effect of the increasing capital resources that are allocated to the agricultural sector. One key concern is that, as more land concentrates in fewer hands, the few large actors can exert power on land markets by dominating prices through regulating supply and demand of land. Unfortunately, empirical evidence on the degree of market power on land markets remains scarce, mainly due to a lack of ownership data. We shed light on the concentration of ownership in agricultural land by analyzing the complete cadastral records of agricultural land at one point in time and for one district in the federal state of Brandenburg. We present the workflow to process the cadastral data for subsequent analysis in GIS and statistical software packages. For our study area, we derive relative and absolute concentration measures for the ownership in agricultural land. Our results suggest high relative concentration on the district level with a Gini coefficient of 0.85. Within the district, we see varying degrees of land concentration, albeit spatial clusters of high and low concentration. Our methodological approach holds great promise because it can be expanded to larger areas and different time periods. However, the cadastral data does not allow to infer on the underlying corporate structures, such as those of large investors who may own several agricultural companies. Such corporative structures may, through their local subsidiaries that could be spatially clustered, exert market power to the detriment of local land supply markets. Additional data and analysis using, for example, registers of company registers, need to be combined with the cadastral data to reveal such structures.

Keywords: Land ownership; land use; agricultural structure; land concentration.

JEL Code: Q15

Published by
DFG Research Unit 2569 FORLand
https://www.forland.hu-berlin.de

Agricultural Land Markets – Efficiency and Regulation

1 Introduction

Concerns over a seemingly increasing concentration in the ownership of agricultural land in fewer hands are growing in many places (Desmarais et al. 2015, Nickerson et al. 2012, van der Ploeg, Franco and Borras 2015), including in Germany (Bunkus and Theesfeld 2018, Forstner et al. 2011). High concentration of land ownership may cause imperfect competition and affect prices for agricultural land, which in turn can cause welfare losses (Cotteleer, Gardebroek and Luijt 2008). However, empirical evidence about the degree of land concentration are scant, mainly because spatial data about ownership can rarely be accessed. In Germany, to the best of our knowledge, an empirical assessment of the spatial concentration of ownership in agricultural land is lacking to date.

Land is an immobile production factor. The concentration of land in a confined region, such as administrative entity, can hence be measured using spatial data. Concentration of land in few hands within the region of interest may result in market power due to the small number of market participants. The market power can be quantified using relative and absolute concentration measures within the region. Relative measures include the Gini coefficient and the Herfindahl-Hirschman index (HHI). Absolute concentration measures can be expressed as the share of farmland owned by a predefined number of the largest owners within the region (Desmarais et al. 2017).

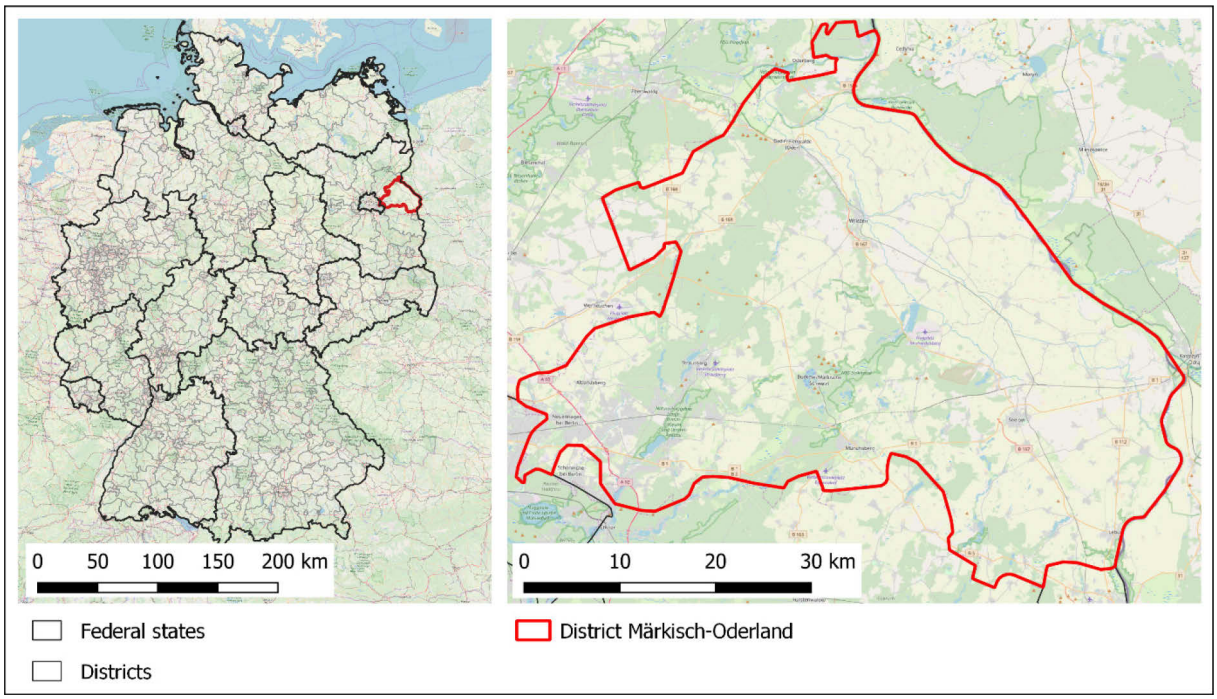
Here we present relative and absolute measures of the concentration of agricultural land in 2019 for the district of Märkisch-Oderland in the federal State of Brandenburg in Germany. We calculate the concentration measures using georeferenced information on land ownership from the official cadastral land register of Germany using the entire cadastral entire for the district. We present descriptive summaries, graphs, and maps of the relative and absolute concentration measures for the entire district, at the level of municipalities (*Gemeinden*), and for local subdistricts (*Gemarkungen*) within Märkisch-Oderland. Overall, we find low relative and absolute concentration of land ownership, except some spatial clusters with higher concentration. Our approach using only the cadastral records must be taken with a grain of salt because it does not reveal higher-level structures, such as if large investors own more than one company in a confined region.

2 Data and methods

2.1 Study area

Märkisch-Oderland covers an area 2,159 km^2, has a population of 193,000, and a population density of 89 person per km^2 (these and the following numbers in this paragraph are for 2016 and were derived from the Landratsamt Märkisch-Oderland 2019). The district borders Poland to the east and Berlin to the west (Figure 1). Agricultural land covers 61% of the district surface and 23% are forested. Of the 126,305 ha of cultivated agricultural area in 2016, 116,972 are arable land and 9,006 ha are grassland. The number of farms in the district decreased from 547 in 2007 to 454 in 2016 while the extent of land remained almost stable; hence average farm size increased from 232 ha to 278 ha. During the same period, the share of organic farms in total farms increased from 7.7% to 10.8%.

Figure 1. Geographic location of the municipality Märkisch-Oderland (Brandenburg) within the Federal Republic of Germany. Background: OpenStreetMap base layer.



2.2 Cadastral data with land ownership

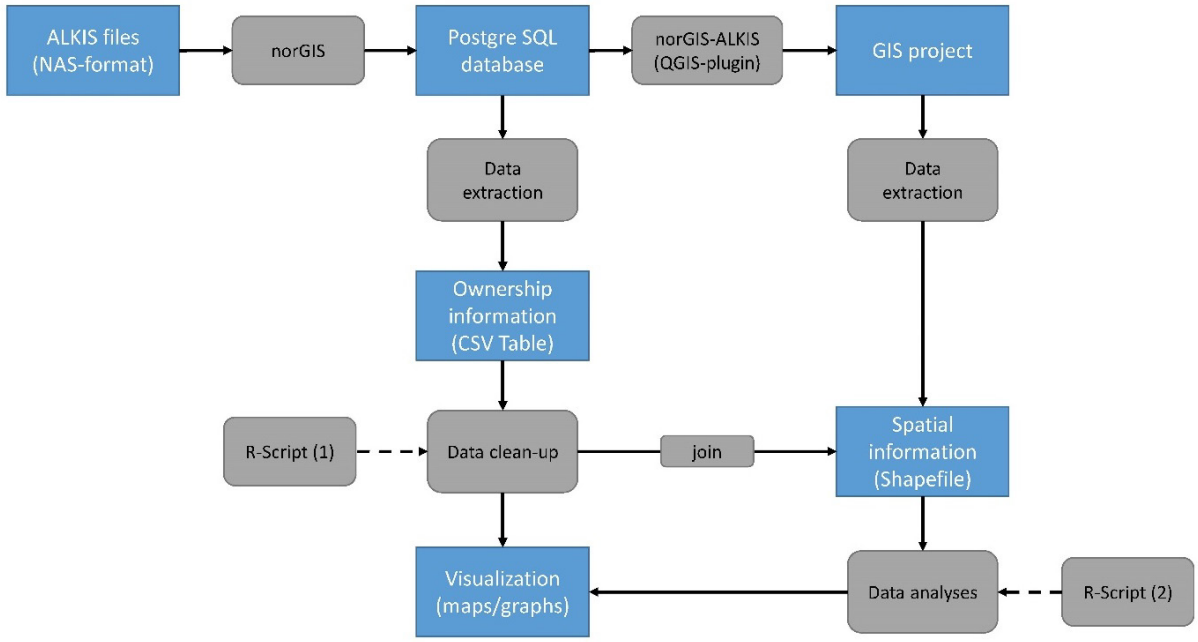
We used the official Information system of the cadastral land register (*Amtliches Liegenschaftskatasterinformationssystem, ALKIS)*. ALKIS combines all real estate information, which were stored either as maps or books prior to the digitalization of the land register. Since 2015, the ALKIS has been introduced to all federal states of Germany with the aim to assure a uniform real estate cadaster without redundancies and based on international norms and standards. Among others, ALKIS holds information on the spatial extent of all properties as well as personal details (e.g., name, address, date of birth) of the individual owners. Therefore, the data are well suited to identify patterns of concentration in agricultural land ownership. We are only away of one other analysis in Germany that used the cadastral data. This application connected the ownership data with taxpayers' records for a municipality in Saxony-Anhalt to assess the share of tax income that is lost to local authorities because of land owners who reside elsewhere (Tietz 2019).

We acquired ALKIS data for the reporting date of February 15, 2019, for all agricultural land parcels of the district *Märkisch-Oderland*. The data were provided in the norm-based data interface (NAS) format by the data provider (the *Landesvermessung und Geobasisinformation Brandenburg, LGB*). The data came as four separate geographical subsets with a total zipped size of 114 MB. The zip-containers contained 196 XML files with a total size of 3.24 GB. To open the files, we created a local PostgreSQL database using PostgreSQL Version 11[1] that served to import the XML files. For the import the norGIS-ALKIS[2] tool was used, which is an open tool designed for the ALKIS NAS-data import into PostgreSQL/PostGIS databases, and based onGDAL/OGR libraries (see a screenshot of the user interface interface Appendix 1; we recommend to follow the installation guide using OSGeo4W). During the import, we ignore all error messages because these were caused by duplicated geometries within the XML files (possibly due to the data provision in several geographical subsets) and would have caused the termination of the data import process. Once the data had been imported, they can be accessed with pgAdmin. To extract the relevant ownership information from the database, we called the table "*v_eigentümer*" (located in Schema → Public → Views) and exported as a CSV table for data clean-up (cf. section 4) and subsequent analyses (cf. section 5). Figure 2

[1] https://www.postgresql.org/download/windows/ last access: 22.07.2019
[2] http://norbit.de/68/

visualizes the entire workflow of extracting the ownership information and parcel boundaries contained from ALKIS into a GIS project.

Figure 2. Workflow of the ALKIS data import into a PostgreSQL database, a QGIS project, and subsequent data clean-up and analyses.



2.3 GIS Integration

We integrated the PostgreSQL database into a Geographic Information System (QGIS Version 3.4.3 Madeira, https://qgis.org/de/site/) using the ALKIS-GIS import tool (http://www.norbit.de/75/, screenshot of the user interface in Appendix 2). We extracted agricultural areas and subsequently the size of the individual parcels using the export data function (screenshot of the GIS project in Appendix 3). The parcel shapefile was then joined with the "*v_eigentümer*" table, which contains information on the landowners, using the field "*gml_id*" (parcel ID). We ensured the persistency of the character encoding to maintain complete and original data entries during the data export steps by maintaining ISO 8859-1 (Latin-1). Joining the tables to the shapefiles with this encoding worked reliably with R (https://www.r-project.org/). The related R-code is available in Appendix 7.

2.4 Data preparation to identify unique ownership features

In a next step, we derived unique identifiers for the individual owners contained in the ALKIS database. The heterogeneity of the proprietary attributes required a multi-step procedure to standardize these complex text attributes. We followed a three-step protocol, which involves 1) categorization of the proprietary type, 2) derivation of a unique proprietary identifier, and 3) text cleaning and standardization.

2.4.1 Categorization of proprietary type

Here we classified individual proprietary information into the following distinct categories:

- Individual private owners or groups of individuals
- Public agencies
- Companies
- Church
- Others (containing, e.g., registered associations (*e.V.*) and foundations)

The shapefile has 74,000 land parcels for Märkisch-Oderland. We relied on the programming environment R to convert the ownership attributes into an object list of type „character string“. The proprietary information in the ALKIS database are diverse in terms of detail, information content, and syntax. We identified and accounted for formal differences in the attribute information to produce a reliable classification of ownership categories.

Below we list anonymized examples of the proprietary attributes of an individual private owner (a), and a private company (b), as they may appear in ALKIS.

a) „Müller, Martin, * 1968-01-31, Hauptstraße 123, 10101 Hauptstadt“

b) „Grüne Bauern GmbH, Landstraße 123, 12345 Wiesenstadt-Auerbach“

A first separation of the proprietary information allowed to separate individual private owners from the remaining ownership types. The information on birth dates, typically introduced through an asterisk (*; see example a) above), facilitated the identification of private owners. Filtering the entries with an asterisk yielded a first rough discrimination of owner categories into individual and non-individuals. The subsequent sanity-checks and further subdivisions were conducted separately for the two resulting datasets.

We subdivided entries with birth dates into segments using the comma separators in a) above, which permit to extract a list of character strings for each entry in the ALKIS database. For the example a) above, this returns:

„Müller“ „Martin“ „ 1968-01-31“ „Hauptstraße 123“ „10101 Hauptstadt“*

The sequence surname(s), prename(s), date of birth, street and number, postal code, and place of residence were consistent for most entries in the database. Prename(s) and surname(s) were given in almost all cases, with few exceptions containing only a surname without additional information. The information on the place of residence was not considered relevant for the generation of a unique proprietary identifier, as potential relocation of an individual can make this attribute ambiguous.

Heterogeneous spelling and multiple notation styles of identical attributes posed challenges to the unique attribution of owners. We therefore applied additional processing and filtering of the character strings. For example, the following notation of the „Grüne Bauern GmbH“ might occur:

„Grüne Bauern GmbH, Landstraße 123, 12345 Wiesenstadt-Auerbach“
„Grüne-Bauern GmbH, Landstr. 123, 12345 Wiesenstadt-Auerbach“
„Gruene Bauern GmbH, Landstr. 123, 12345 Wiesenstadt (Auerbach)“

The generation of a unique attribute was thus primarily based on sur- and prename(s). A sanity check on 10,000 entries revealed that some attribute fields of companies contained asterisks, which required an additional rule. We hence add an additional subdivision based on the observation that the first comma-separated attribute of individual private owners (i.e., the surname) contained only a single word in almost all cases (e.g., Müller) and hyphens commonly combined double-surnames (e.g., Müller-Bauermann). On the contrary, public agencies and companies commonly contained multiple words, separated by spaces. Consequently, we further subdivided the surname attribute with space characters to classify all entries with only one word as private owners. In that way, public agencies, companies, church, and other types of owners could be identified under the first condition of not having an asterisk in the attribute field and the second condition that the first comma-separated section contained more than one word.

The remaining cases consisted of individual owners with an asterisk (birthdate) and two words in the first comma-separated attribute, or those without asterisk and only one word in the first comma-separated attribute. Based on a manual check of these cases, several lists of keywords were created to classify the uncertain cases as either a public agency, company, church, or other type of proprietary:

"AG", "Agrar", "Agraraktiengesellschaft", "Agrargenossenschaft", "Agrargesellschaft", "Agrarproduktion", "Agrarwirtschaftsgesellschaft", "Agro", "Agrogenossenschaft", "AGW", "Anlieger", "Bauernsiedlung", "Bodenfonds", "BRD", "Büdnergenossenschaft", "Bundesanstalt", "Bundesfinanzverwaltung", "Bundesrepublik", "bürgerlichen", "BVVG", "Deutschland", "Erben", "Evangelische", "Familiengesellschaft", "Fischergemeinde", "Forstliegenschaft", "Forstverwaltung", "GbR", "Gebrüder", "Gemeinde", "Gemeinde", "Gesellschaft", "Gestüt", "GmbH", "Grundstücksfonds", "Haus-", "Herrenlos", "interAconsult.", "Kirche", "Kirchengemeinde", "Kossätengenossenschaft", "Land", "Landesvermessung", "Landfarm", "Landwirtschaft", "Landwirtschaftliche", "LPG", "NOGA", "ODEGA", "Pfarre", "Produktionsgemeinschaft", "Produktionsgenossenschaft", "Separationsinteressenten", "Stephanus-Stiftung", "Treuhandanstalt", "Unbekannte", "Vermögensverwaltungsgesellschaft".

Finally, we manually cleaned the resulting entries. For example, we shifted the *Bodenverwertungs- und -verwaltungs GmbH (BVVG)* to the public agencies because the BVVG is a state company with the original mission to privatize formerly collectively and state-owned agricultural and forested areas. We also shifted the Nature and Biodiversity Conservation Union (*Naturschutzbund Deutschland e.V., NABU*) to the category “others”.

2.4.2 Unique proprietary identifier

We created an unambiguous proprietary identifier based on the attributes’ pre- and surname of individual private owners or based on the attribute name of public agencies, companies, church, or other types of owners.

2.4.3 Text homogenization and cleaning

Ownership information contained nuanced variation in the writing and notation style, which complicates the unique attribution of individual owners. For instance, different spelling of property entries pose challenges, such as „Grüne Bauern GmbH“, „Grüne-Bauern GmbH“, or „Die Grüne Bauern (GmbH)“.

We performed further text homogenization and cleaning to counteract these issues. We removed all punctuation and special symbols (including, for instance, line breaks, spaces, hyphens, quotes, semicolons, colons, and “&”). We also deleted small connecting words, such as *in, der, die, den, dem, des, bei, von, zu, und*. Finally, we removed all capitalization and converted German umlauts (a, ö, ü) into vowels (ae, oe, ue). This procedure resulted in unambiguous and unique identifiers, such as “gruenebauerngmbh”.

Based on additional keyword lists, we further separated public agencies, companies, church, and other types of owners. Public agencies contained one of the words *' land ', 'gemeinde ', 'stadt ', 'bundes ', 'bundesrepublik', 'landes ', 'eigentumdesvolkes ', 'treuhandanstalt '*. Companies contained *'gmbh', 'mbh', ' ag ', ' ag$', ' gbr ', ' gbr$', 'genossenschaft', 'aktiengesellschaft', 'agrargenossenschaft', 'agrargesellschaft', ' fonds ', 'stiftung*', ' eg ', ' eg$', ' kg ', ' kg$', ' ug ', ' ug$', 'vkg', ' lpg', ' ohg'*. The occurrence of the words *'kirche ', 'evangelisch', 'katholisch', ' pfarre'* indicate ownership of the church. Keywords, such as *' ev ', ' ev$',* and *'unbekannt'* identified other types of owners.

2.5 Quantification of ownership concentration

To quantify concentration of ownership of agricultural land in Märkisch-Oderland we ranked the largest owners in terms of area, which we further subdivided into the distinct owner categories, such as private individuals, companies, or public agencies. In addition, we calculated several relative and absolute concentration metrics.

The Lorenz curve, also known as the disparity curve, allows a graphic representation of the distribution of the data (here ownership) and shows any disparity (inequality) or concentration (equality) that exists. The data are sorted in ascending order, summed up, and displayed in a diagram. The proportions of the total of the analyzed feature (here owners) are plotted on the x-axis, and the proportions of the feature sum (here areas) appear on the y-axis. The resulting curve shows the shares of the total sum that account for the different shares of the basic

quantity of the feature. The shape of the curve below the 45-degree line allows to interpret the inequality in the distribution, where the 45-degree line corresponds to a perfect uniform distribution where every owner within the confined boundary owns exactly the same amount of land.

From the Lorenz curve, the so-called Gini coefficient can be calculated as a disparity measure, with values ranging from zero to one (the higher the coefficient, the more unequal the distribution). The Gini coefficient is calculated from the area fraction between the 45-degree line (perfect distribution) and the Lorenz curve. Consequently, the Gini coefficient increases with a higher concentration.

The Herfindahl-Hirschman index (HHI) is a relative concentration measure whose value increases at a higher concentration. The HHI is formed from the sum of all squared market shares of the competitors of a market and is, depending on the scale, between 1 / N and 1, where 1 corresponds to a monopoly.

3 Results

3.1 District (Landkreis)

We calculated the concentration patterns from ALKIS for all agricultural land parcels in Märkisch-Oderland using the PostgreSQL database and the PostGIS interface (described above in Figure 2). Based on the cleaned-up dataset, we identified a total of 16,083 owner entities with 74,690 parcels totaling in 143,729 ha (Table 1) were identified for the district (note that this area is lower than what is reported as agricultural land by the 126,305 ha of cultivated agricultural area reported by the district authorities, arguably because not all land parcels that appear in the cadastral records are cultivated). The *Bodenverwertungs– und verwaltungs GmbH (BVVG)* is the largest single landowner with an area of 5,346 ha (3,7% of all agricultural land in the district).

Table 1. Descriptive statistics of the agricultural parcels contained in the ALKIS data.

Minimum [ha]	Mean [ha]	Maximum [ha]	Total [ha]
<0.001	1.55	143.13	143,729

To examine the concentration of ownership, we analyzed the unique entries of owner characteristics, which we separated into the five categories “Individual private ownership”, “Public agencies”, “Companies”, “Church”, and “Others” (Table 1, Figure 3). The class “Other” contains a few parcels for which attributes were erroneous, and parcels that belong to the environmental association Nature and Biodiversity Conservation Union (NABU). Overall, this class contains only 62 parcels with a total size of 1,529 ha (Table 2). The main share of agricultural land (88,278 ha or 61%) belongs to 14,877 private individuals who own 5.9 ha on average. Parcels owned by private individuals may also belong to more than one unique owner, such as in case of joint heirship (*Erbengemeinschaft*). In these cases, we only considered the first data entry. The 561 private companies own 24% or 35,159 ha of the agricultural land in the district (on average 62 ha per company) (Table 2). With reference to all unique owner entries and the associated areas, the overall Gini index for Märkisch-Oderland is 0.851, calculated from the associated Lorenz curve (Figure 4), which suggests high ownership concentration of agricultural land for the entire district.

The spatial patterns that result from the aggregation of the parcels into the categories are shown in Figure 5 for four selected areas within Märkisch-Oderland.

Figure 3. Agricultural land and number of individual parcels for the five ownership classes.

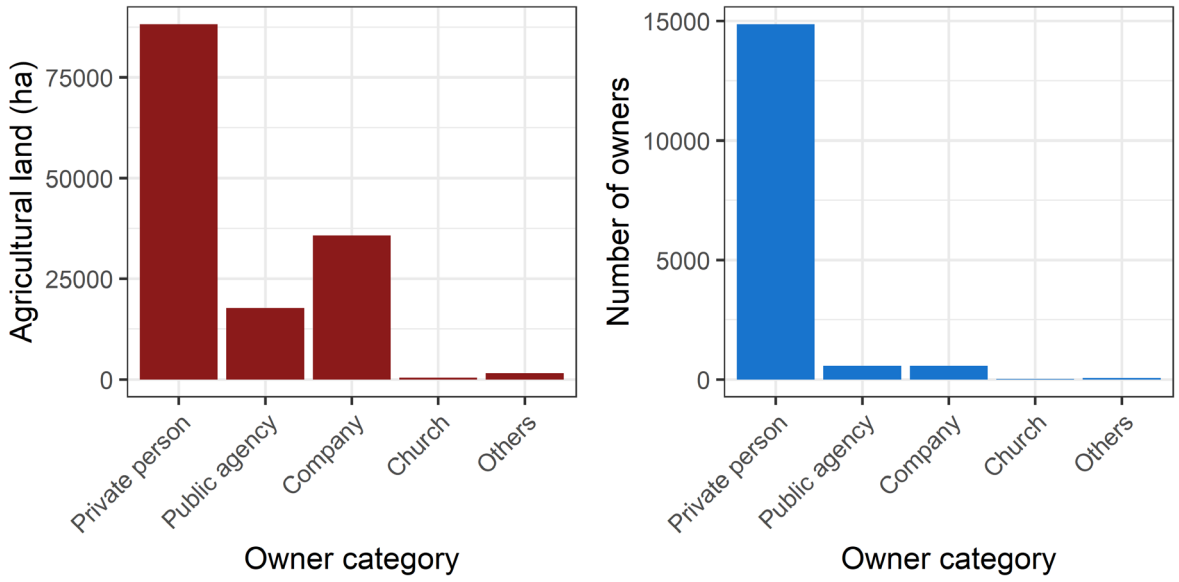


Table 2. Shares of agricultural areas of private individual owners and private or public entities in Märkisch-Oderland.

Category	Number of owners	Mean area per owner [ha]	Mean number of parcels [n]	Total area owned [ha]
Private individual owners	14,877	5.9	3.1	88,277.9
Public agencies	565	22.2	17.7	12,526.3
Companies	583	71.6	30.9	41,726.6
Church	15	29.2	10.4	438.2
Others	62	12.3	6.6	760.5
Overall	16,095	8.9	4.6	143,729.4

Figure 4. Lorenz curve with Gini coefficient and Herfindahl-Hirschman index (HHI) for relative ownership concentration of all agricultural areas in Märkisch-Oderland.

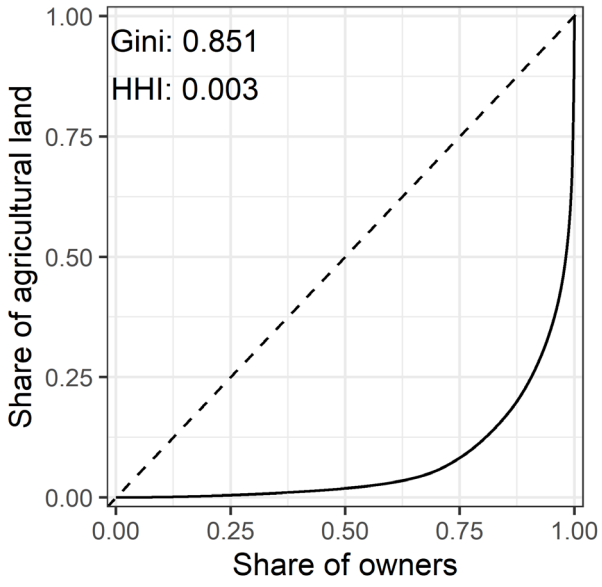
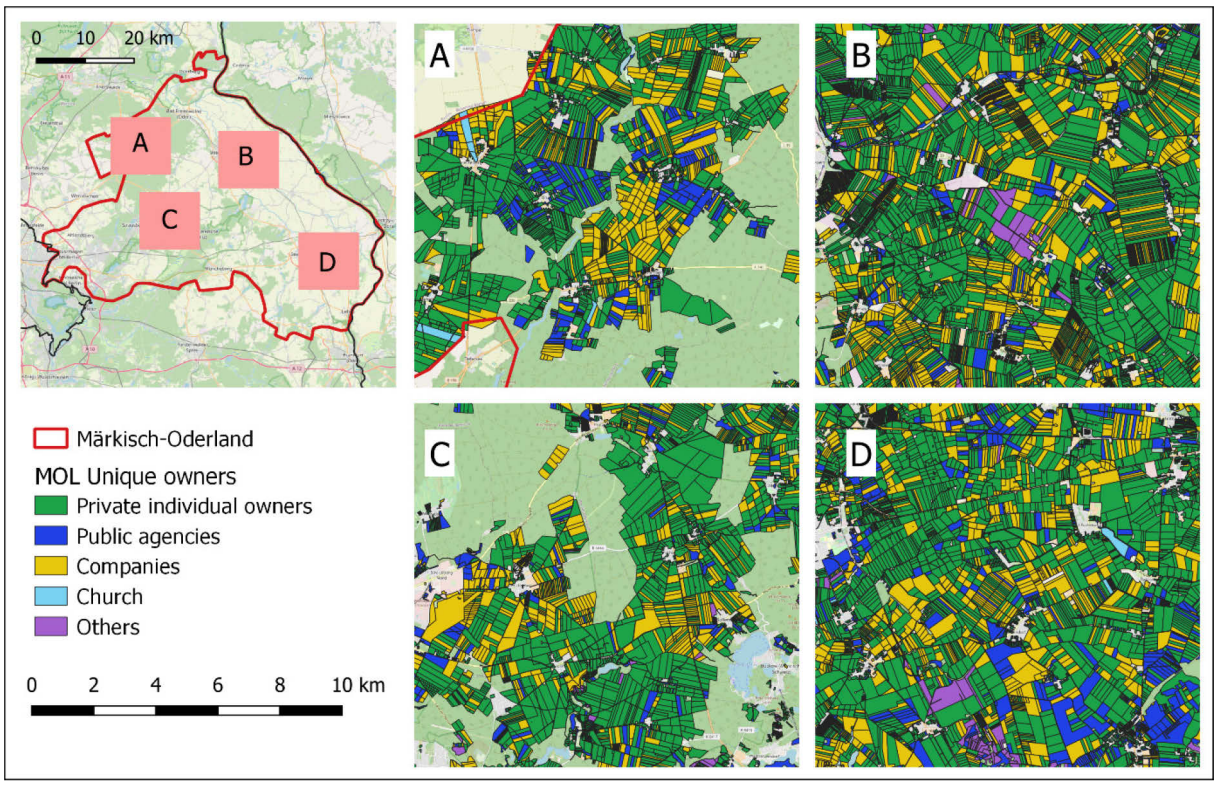


Figure 5. Examples for the distribution of parcels with unique owners in Märkisch-Oderland (MOL). The different colors highlight to which group of owners the parcels belong to. A, B, C, and D zoom into the spatial patters of ownership patterns.



3.2 Municipality *(Gemeinde)*

We calculated the concentration measures for each of the 45 municipalities in Märkisch-Oderland. The Gini-coefficient ranges between 0.72 and 0.87 (mean: 0.8; median: 0.81). The respective Lorenz curves for two selected municipalities are shown in the Figure 6. A summary of all Gini coefficients for each municipality can be found in Appendix 4 and the spatial patterns of the Gini coefficients are mapped in Figure 7. Appendix 5 shows the number of owners per category and municipality and Appendix 6 depicts the area and the share in total area of the largest owners per municipality and category.

Figure 6. Lorenz curves for two municipalities in Märkisch-Oderland with the lowest (Reitwein, left) and highest (Prötzel, right) Gini coefficients, including the Herfindahl-Hirschman indices (HHI).

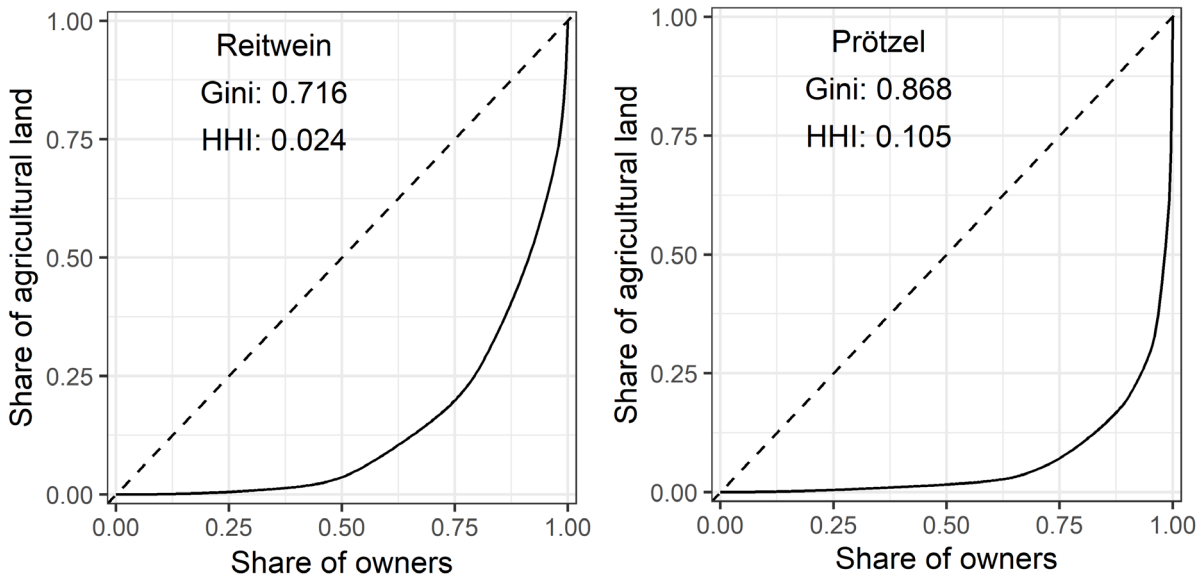
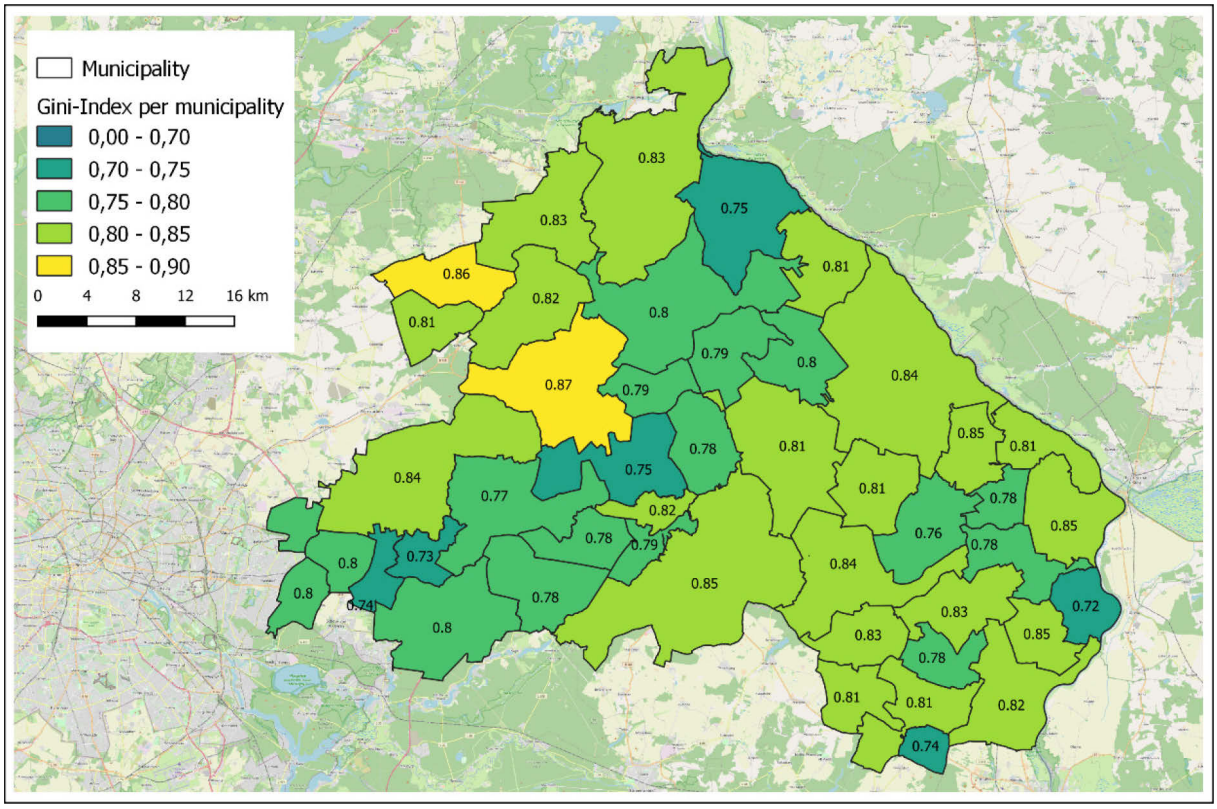
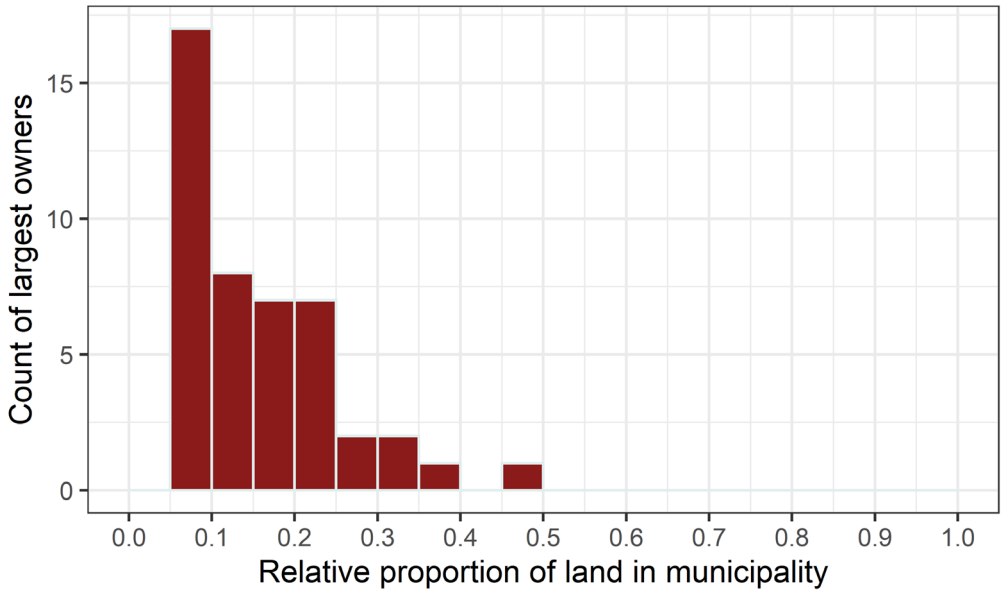


Figure 7. Spatial pattern of municipality-level Gini coefficients for Märkisch-Oderland.



Most Gini coefficients on municipality level are above 0.8, suggesting a high ownership concentration in Märkisch-Oderland (Figure 7). However, the largest landowners in each of the 45 municipalities in Märkisch-Oderland rarely own more than 50% of the land within a municipality. This is illustrated in Figure 8, which shows the number of the largest landowners (y-axis), along with the relative proportion of the land each of them owns within the respective municipality (x-axis). Figure 8 corroborates that the largest landowners rarely hold more than 50% of the agricultural areas within one municipality. The high Gini coefficients on municipality level resulted mainly from the large number of individual landowners that own comparably small shares of land.

Figure 8. Histogram of the share of land owned by the largest landowner in a municipality (n=45).



3.3 Subdistrict (Gemarkung)

We quantified concentration for all 183 subdistricts of Märkisch-Oderland that contain agricultural areas. The number of owners within each subdistrict ranges between 1 and 566 with a mean number of parcels per owner ranging from one to ten. This results in Gini coefficients between 0 and 0.89 (mean: 0.72; median: 0.75). Figure 9 visualizes the Lorenz curves for a subdistrict with a low Gini coefficient (left) and the highest Gini coefficient (right). The coefficients of 0 (n=4) stem from subdistricts with an agricultural area in the cadastral records of less than 5.1 ha, which is owned by a single owner. At the administrative level of subdistricts, the ownership concentration tends to be lower than at the municipality level (compare Figure 7 and Figure 10).

Figure 9. Example of Lorenz curves for two subdistricts, one with a low Gini (Sonnenburg, left) and one with the highest Gini coefficient (Diedersdorf, right), HHI: Herfindahl-Hirschman indices.

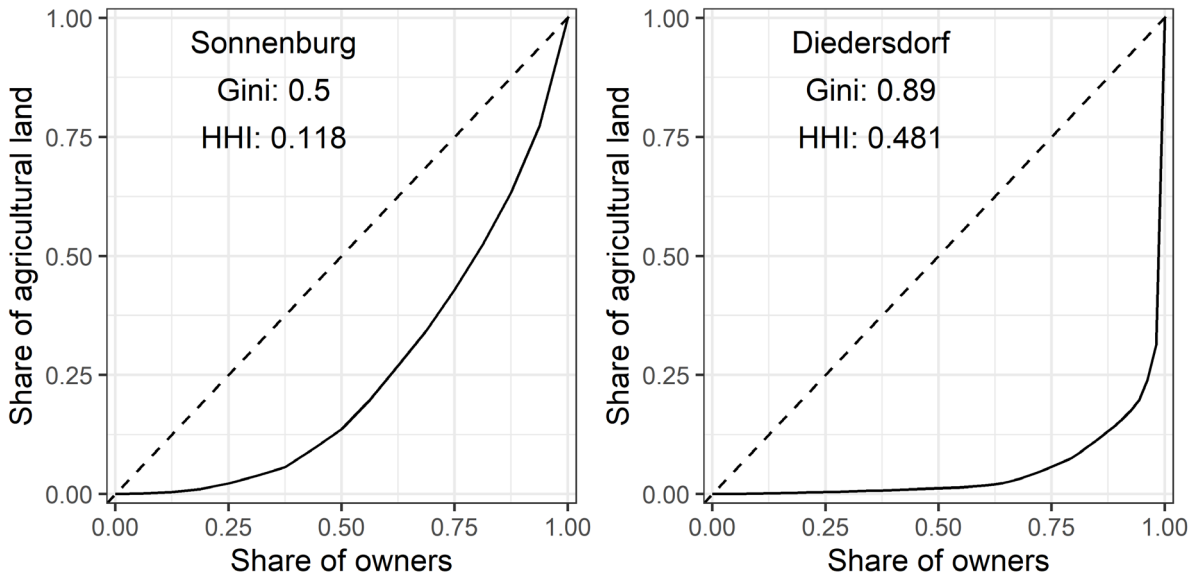
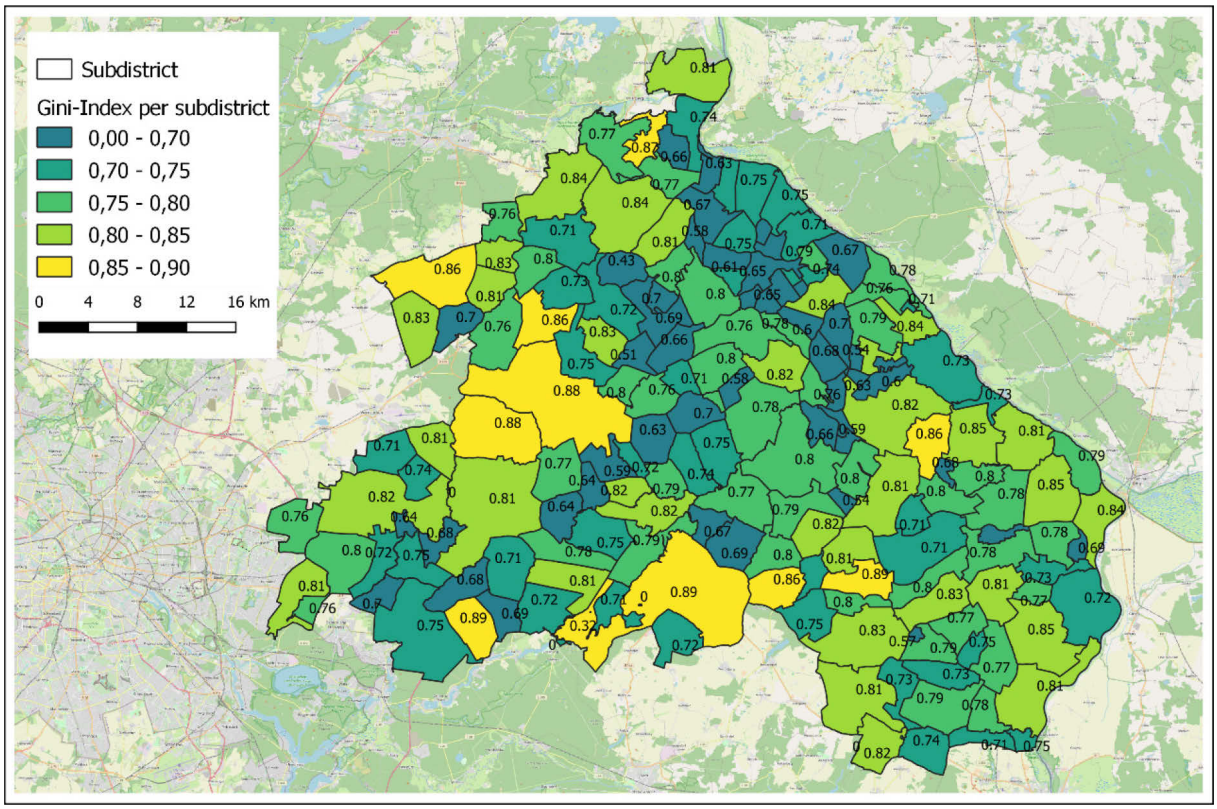


Figure 10. Spatial pattern of the Gini index on subdistrict level for Märkisch-Oderland.

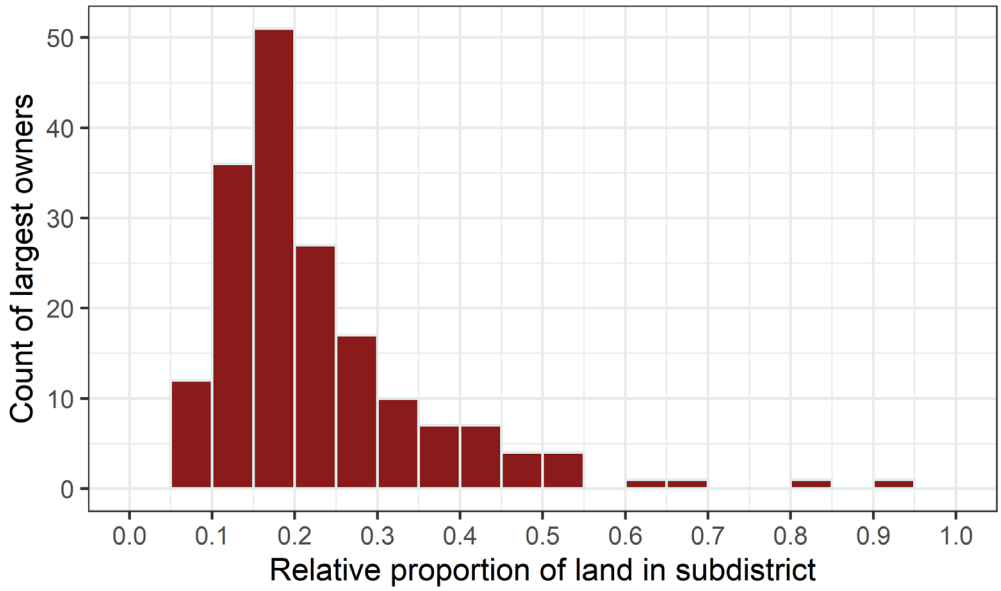


Similar to the municipality level, we find only eight subdistricts where the largest landowners hold more than 50% of the agricultural land within the subdistrict (Figure 11, Table 3). Relative concentration tends to be lower in smaller subdistricts (Figure 10).

Table 3. The eight subdistricts where the largest owners possess more than 50% of agricultural land within the subdistrict.

Subdistrict	Category	Area [ha]	Area [%]
Müncheberg	Private individual owners	2	92.5
Herzfelde	Private individual owners	328	53.9
Wesendahl	Private individual owners	483	53.7
Blanke Heide	Companies	24	84.1
Diedersdorf	Companies	365	68.7
Alt Rosenthal	Companies	289	60.1
Bollersdorf	Companies	246	52.2
Harnekop	Companies	273	50.5

Figure 11. Histogram of the share of land per subdistrict owned by its largest landowner (n=179).



4 Conclusions

We provided an assessment of the concentration of agricultural land using official cadastral data for the district of Märkisch-Oderland in Northeast Germany. We presented the workflow to extract the data from the cadastral system with the aim to allow replication of extracting ownership records and concentration from the same data. Considering that more statistical data will be accessible for free in Germany, following a recent EU directive, we anticipate that cadastral records will become available in more federal states, as they have already in Brandenburg, albeit without the detailed ownership information to protect the privacy of owners.

The analysis of the cadastral ownership records revealed modest to high degrees of relative concentration of agricultural land with an average Gini coefficient of 0.85 on district level, albeit with a large spatial variation within the district. We find cases where companies and, to lesser extent, individuals possess a large share of land within a region that may cause some degree of market power on local land markets. The high degrees of relative ownership concentration partly arise because many landholders own small parcels of land within an administrative region.

It is important to note that the ownership data does not account for corporate structures. For example, a larger investor may possess several subsidiary companies that operate within a smaller region. In such cases, the factual land concentration can be substantially larger than estimated by our approach. The ALKIS data do not permit to reveal such business structures but they could help to pinpoint locations where, for example, a local survey of land holding would allow to disclose such corporate structures. Another option is to connect the cadastral records with register of companies (*Handelsregister*) to reveal such ownerships structures (see Trautvetter and Henn, 2020, for an example of ownership of real estate in Berlin).

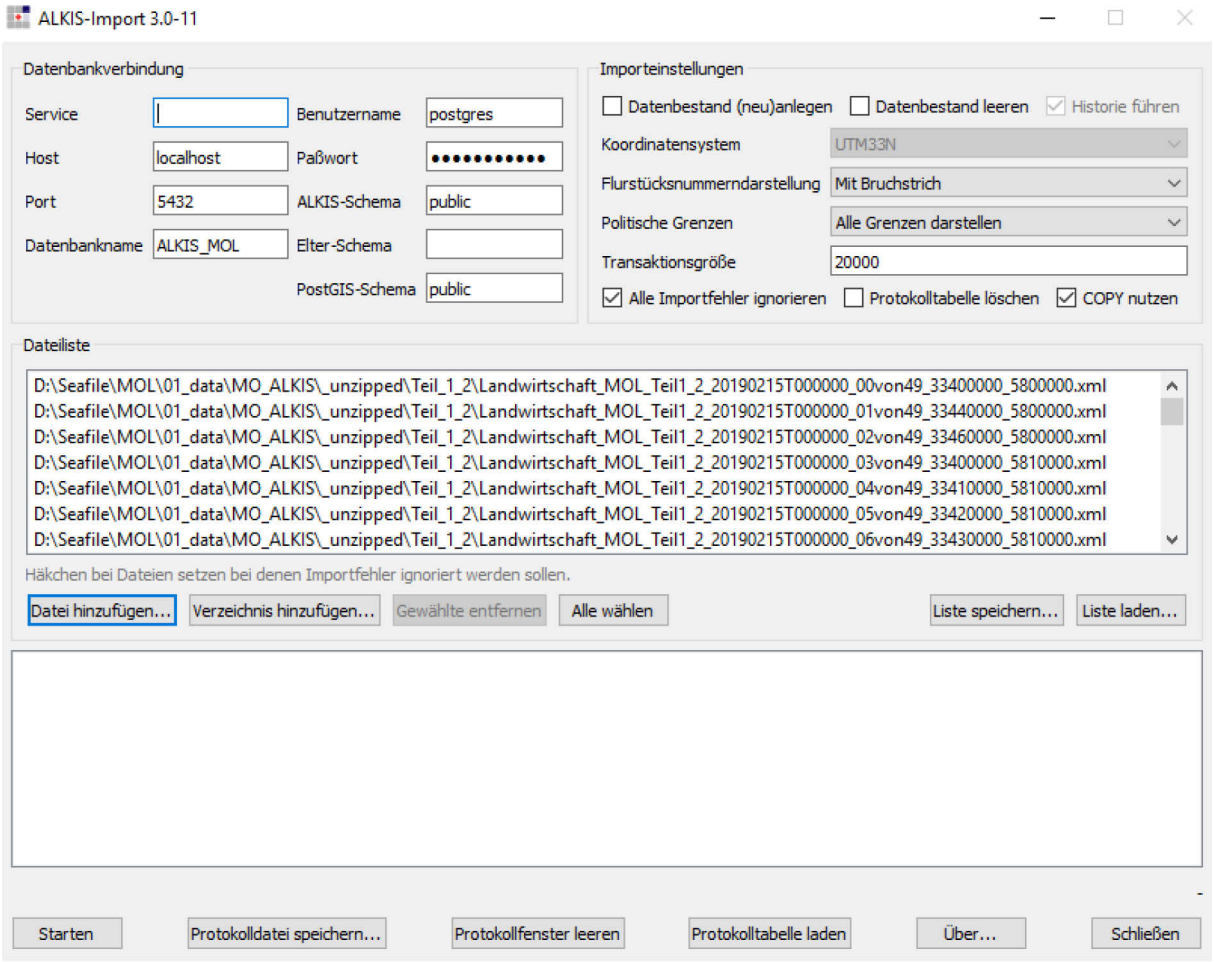
Our work paves the way for several useful extensions. First, we present a procedure that is, to a large extent, automated and may hence be followed with other ALKIS data, such as for all of Brandenburg or, if computational resources allow, beyond. Another attractive application using the ALKIS data would be to analyze different time slices, provided ownership data for earlier points in time will be available and accessible in similar quality and depth. Analyzing changes over time may reveal if concentration on land markets has indeed increased, as many grey literature sources and media reports suggest.

5 Glossary

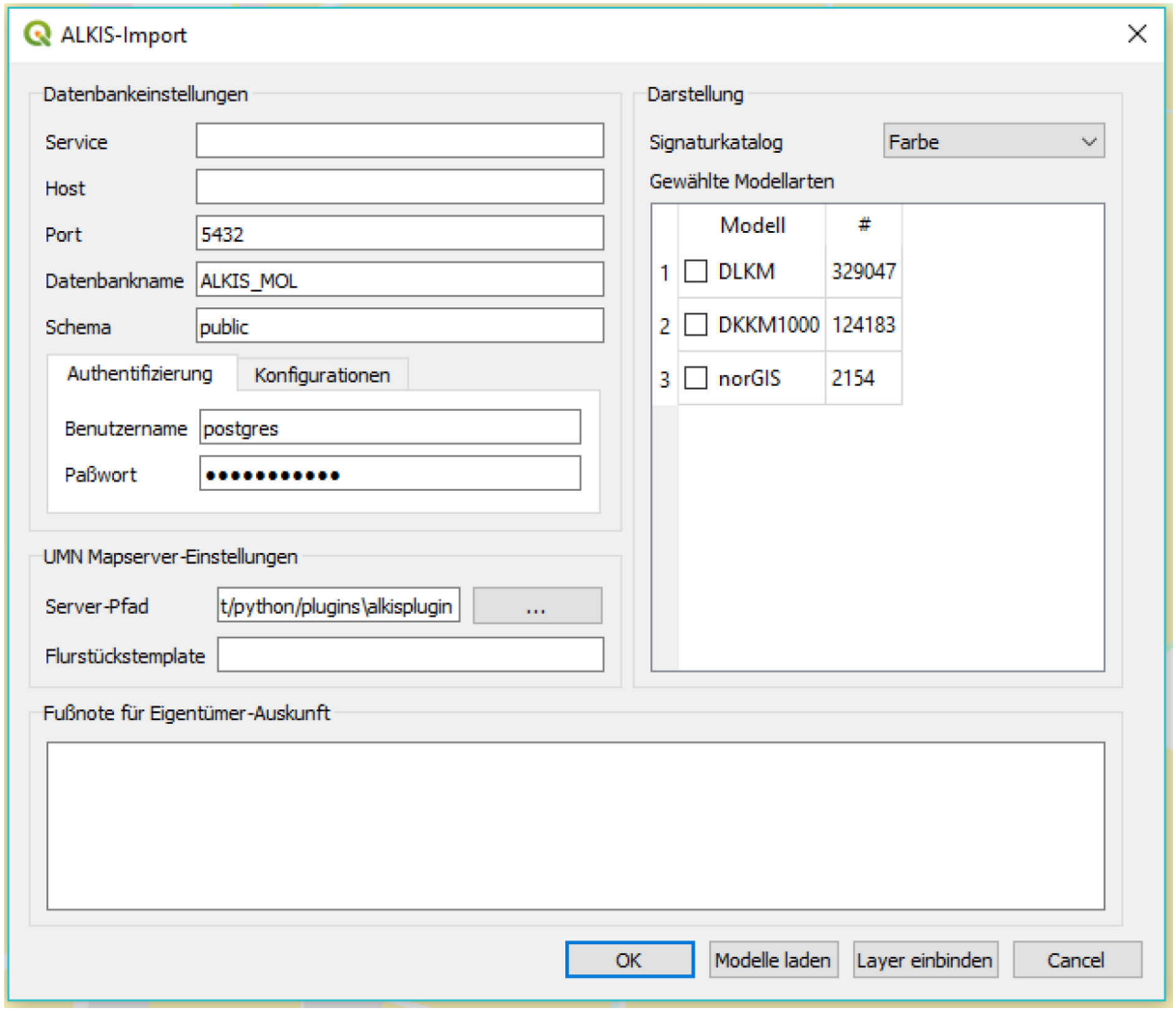
Landkreis	District
Gemeinde	Municipality
Gemarkung	Subdistrict
Flurstück	Land parcel
Eigentümer	Owner (*de jure*)
Privatwirtschaftliches Unternehmen	Private company
Staatliche Gesellschaften	Public agency
Privatperson	Private individual

6 Appendices

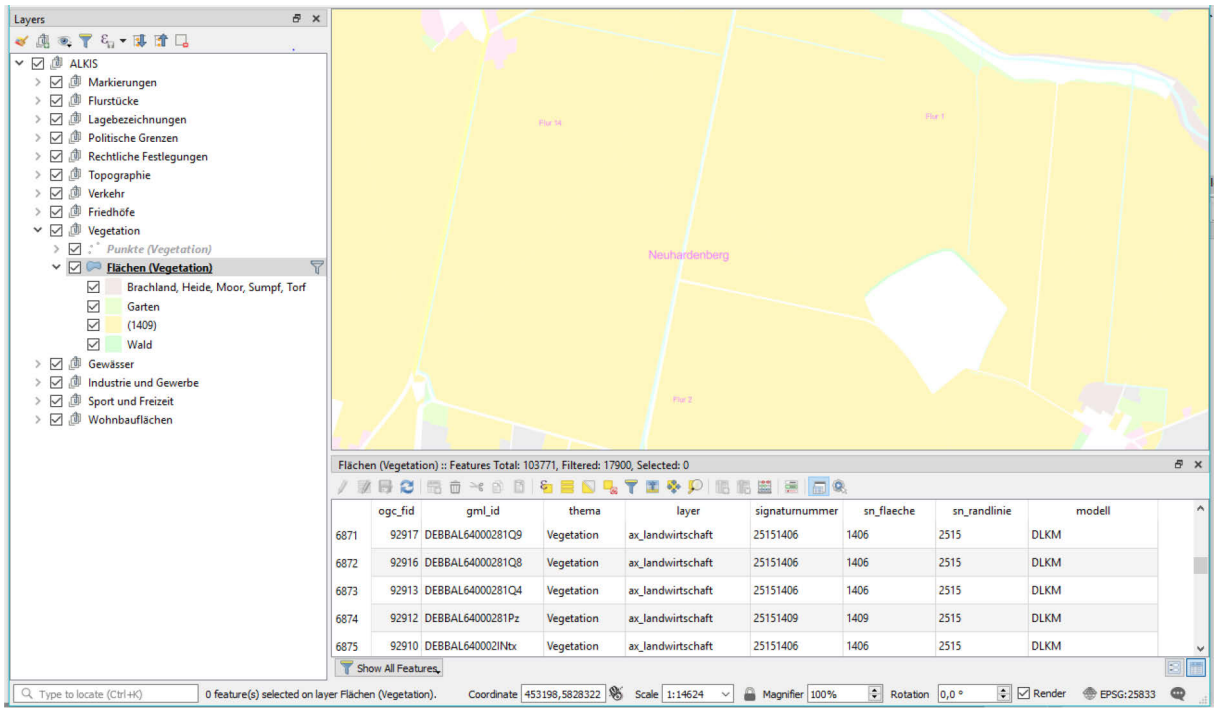
Appendix 1. norGIS ALKIS data import widget.



Appendix 2. norGIS ALKIS import QGIS plugin.



Appendix 3. Screenshot of the ALKIS QGIS project for Märkisch-Oderland.



Appendix 4. Area, parcels size, number of parcels, and Gini coefficients per municipality.

Municipality name	Total area [ha]	Mean parcel size [ha]	No. of parcels	Gini
Alt Tucheband	2,773.33	2.66	364	0.78
Altlandsberg	6,097.85	5.05	632	0.84
Bad Freienwalde (Oder)	6,046.73	4.22	1,521	0.83
Beiersdorf-Freudenberg	1,846.84	3.27	169	0.81
Bleyen-Genschmar	2,731.76	4.02	320	0.81
Bliesdorf	2,843.27	5.43	276	0.79
Buckow (Märkische Schweiz)	393.08	2.86	116	0.82
Falkenberg	3,514.95	4.58	590	0.83
Falkenhagen (Mark)	1,689.80	3.86	168	0.81
Fichtenhöhe	2,172.73	3.50	199	0.78
Fredersdorf-Vogelsdorf	528.39	3.12	218	0.74
Garzau-Garzin	1,713.75	4.59	186	0.78
Golzow	1,598.42	3.43	219	0.78
Gusow-Platkow	2,731.83	3.48	525	0.81
Heckelberg-Brunow	2,413.04	4.66	142	0.86
Höhenland	2,648.48	3.45	286	0.82
Hoppegarten	1,482.78	3.32	300	0.8
Küstriner Vorland	4,064.08	3.29	545	0.85
Lebus	4,838.19	3.88	671	0.82
Letschin	13,212.70	4.44	1,490	0.84
Lietzen	1,809.41	4.72	114	0.83
Lindendorf	3,722.11	3.36	406	0.83
Märkische Höhe	2,344.10	5.12	187	0.78
Müncheberg	8,766.48	3.30	1,090	0.85
Neuenhagen bei Berlin	764.41	3.85	75	0.8
Neuhardenberg	4,538.94	4.84	461	0.81
Neulewin	3,727.16	4.82	462	0.81
Neutrebbin	3,420.64	4.09	433	0.8
Oberbarnim	3,576.44	3.69	399	0.75
Oderaue	6,024.83	5.90	752	0.75
Petershagen/Eggersdorf	352.58	3.33	144	0.73
Podelzig	2,326.45	3.12	274	0.85
Prötzel	2,729.62	3.75	249	0.87
Rehfelde	3,539.73	4.51	462	0.78
Reichenow-Möglin	1,942.01	4.40	179	0.79
Reitwein	2,100.00	3.84	242	0.72
Rüdersdorf bei Berlin	2,618.42	3.59	379	0.8
Seelow	3,654.62	4.44	515	0.76
Strausberg	2,331.00	3.86	306	0.77
Treplin	733.49	3.88	73	0.74
Vierlinden	4,939.73	5.35	366	0.84
Waldsieversdorf	320.03	3.67	27	0.79
Wriezen	6,814.66	4.10	882	0.8
Zechin	2,622.88	4.39	407	0.85
Zeschdorf	2,662.57	4.10	268	0.81

Appendix 5. Number of owners per category and municipality.

Municipality name	Private	Public	Company	Church	Other	Total
Alt Tucheband	334	14	15	1	0	364
Altlandsberg	552	41	37	1	1	632
Bad Freienwalde (Oder)	1,399	74	41	2	5	1,521
Beiersdorf-Freudenberg	137	15	14	1	2	169
Bleyen-Genschmar	279	28	13	0	0	320
Bliesdorf	235	15	25	0	1	276
Buckow (Märkische Schweiz)	106	7	3	0	0	116
Falkenberg	523	38	27	2	0	590
Falkenhagen (Mark)	143	13	11	0	1	168
Fichtenhöhe	164	16	17	1	1	199
Fredersdorf-Vogelsdorf	173	18	25	1	1	218
Garzau-Garzin	157	16	11	0	2	186
Golzow	192	10	14	0	3	219
Gusow-Platkow	480	18	21	1	5	525
Heckelberg-Brunow	122	8	11	1	0	142
Höhenland	245	17	22	1	1	286
Hoppegarten	238	30	32	0	0	300
Küstriner Vorland	489	33	22	1	0	545
Lebus	589	45	36	1	0	671
Letschin	1,372	56	58	0	4	1,490
Lietzen	91	9	13	0	1	114
Lindendorf	349	21	33	0	3	406
Märkische Höhe	156	15	15	0	1	187
Müncheberg	996	42	46	0	6	1,090
Neuenhagen bei Berlin	43	14	17	0	1	75
Neuhardenberg	408	17	32	1	3	461
Neulewin	408	28	20	1	5	462
Neutrebbin	383	20	29	0	1	433
Oberbarnim	354	20	22	1	2	399
Oderaue	684	37	24	4	3	752
Petershagen/Eggersdorf	129	7	7	0	1	144
Podelzig	246	12	16	0	0	274
Prötzel	211	15	20	0	3	249
Rehfelde	404	24	31	1	2	462
Reichenow-Möglin	159	10	10	0	0	179
Reitwein	201	23	14	1	3	242
Rüdersdorf bei Berlin	297	42	34	1	5	379
Seelow	448	26	37	1	3	515
Strausberg	251	24	28	0	3	306
Treplin	61	6	5	1	0	73
Vierlinden	317	14	32	2	1	366
Waldsieversdorf	16	7	2	0	2	27
Wriezen	801	35	42	2	2	882
Zechin	374	18	15	0	0	407
Zeschdorf	224	20	23	1	0	268

Note: Owners can possess parcels in more than one municipality; hence the total number of owners here exceeds the total in Table 2.

Appendix 6. Area and area share of the largest owners per municipality and category.

Municipality	Category	Area [ha]	Area [%]
Heckelberg-Brunow	Private individual owners	734	30.41
Prötzel	Private individual owners	785	28.77
Buckow (Märkische Schweiz)	Private individual owners	84	21.48
Treplin	Private individual owners	150	20.39
Falkenberg	Private individual owners	567	16.13
Rüdersdorf bei Berlin	Private individual owners	401	15.32
Bliesdorf	Private individual owners	348	12.23
Altlandsberg	Private individual owners	494	8.10
Petershagen/Eggersdorf	Private individual owners	23	6.52
Waldsieversdorf	Public agencies	145	45.41
Podelzig	Public agencies	488	21.00
Küstriner Vorland	Public agencies	540	13.28
Bleyen-Genschmar	Public agencies	320	11.73
Reitwein	Public agencies	215	10.24
Strausberg	Public agencies	189	8.10
Lietzen	Companies	696	38.45
Falkenhagen (Mark)	Companies	536	31.73
Höhenland	Companies	755	28.50
Beiersdorf-Freudenberg	Companies	436	23.59
Neuenhagen bei Berlin	Companies	169	22.16
Zechin	Companies	566	21.57
Reichenow-Möglin	Companies	414	21.34
Märkische Höhe	Companies	448	19.12
Neulewin	Companies	701	18.80
Fichtenhöhe	Companies	385	17.70
Alt Tucheband	Companies	467	16.83
Fredersdorf-Vogelsdorf	Companies	80	15.11
Seelow	Companies	449	12.28
Lindendorf	Companies	437	11.74
Oberbarnim	Companies	417	11.66
Zeschdorf	Companies	282	10.61
Lebus	Companies	458	9.47
Vierlinden	Companies	466	9.43
Neuhardenberg	Companies	423	9.31
Müncheberg	Companies	787	8.98
Golzow	Companies	142	8.88
Rehfelde	Companies	309	8.74
Garzau-Garzin	Companies	138	8.06
Gusow-Platkow	Companies	215	7.88
Oderaue	Companies	437	7.25
Neutrebbin	Companies	221	6.45
Hoppegarten	Companies	86	5.81
Wriezen	Companies	390	5.72
Letschin	Companies	704	5.33
Bad Freienwalde (Oder)	Companies	314	5.20

Appendix 7. R-code to clean the ALKIS agricultural ownership dataset and extract information for further analyses.

```
### helper function
`%!in%` <- Negate(`%in%`)

### function definition
code_eigentumer_groups <- function(shp, blacklist, public, private, church, unclear){

 ### calculate area based on raster area fct
 shp@data$hectares <- area(shp)/10000

 ### get attribute table
 data <- shp@data

 ### add fid
 data$fid <- c(1:length(data$eigntmr))

 ### add eigent class attribute
 data$prp <- NA

 ### clean text fields
 data$names <- data$eigntmr
 data$names <- gsub('\n', ' ', data$names)
 data$names <- gsub('-', ' ', data$names)
 data$names <- gsub('\"', '', data$names)
 data$names <- gsub('&', '', data$names)
 data$names <- gsub(';', '', data$names)
 data$names <- gsub(':', '', data$names)
 data$names <- gsub('[.]', '', data$names)
 data$names <- tolower(data$names)

 data$names <- gsub('Ã¤', 'ae', data$names)
 data$names <- gsub('Ã¶', 'oe', data$names)
 data$names <- gsub('Ã¼', 'ue', data$names)
 data$names <- gsub('Ã', 'ss', data$names)
 data$names <- gsub(' in ', ' ', data$names)
 data$names <- gsub(' die ', ' ', data$names)
 data$names <- gsub(' der ', ' ', data$names)
 data$names <- gsub(' den ', ' ', data$names)
 data$names <- gsub(' dem ', ' ', data$names)
 data$names <- gsub(' des ', ' ', data$names)
 data$names <- gsub(' bei ', ' ', data$names)
 data$names <- gsub(' von ', ' ', data$names)
 data$names <- gsub(' zu ', ' ', data$names)
 data$names <- gsub(' und ', ' ', data$names)

 ### add placeholder string for fields with no comma.
 data$names <- paste0(data$names, ', xxx.')

 ##############################################################################
 ### first check: separate people & entities based on * initiating birth date
 ##############################################################################

 ### get fid of entries with * (birth date)
 people_index <- data$fid[grep('[*]', data$names)]

 ### get fid of entries without * (birth date)
 entity_index <- data$fid[data$fid %!in% people_index]

 ##############################################################################
 ### check people based on name structure
```

```
################################################################################

### private peoples names, split by comma
people_split <- strsplit(as.character(data$names[data$fid %in% people_index]), ',')

### get last name
people_fst <- lapply(people_split, '[[', 1)

### use first entity name, split by space to select those with more than one entry until comma
### to get entities, falsely allocated to people (usually, entities have several words)
people_fst_split <- strsplit(as.character(people_fst), ' ')

### get fid of people, where we are relatively certain
people_prp <- people_index[lengths(people_fst_split)==1]

### assign property code to people_prp
data$prp[data$fid %in% people_prp] <- 1

################################################################################
### check entities based on name structure
################################################################################

### entity name, split by comma
entity_split <- strsplit(as.character(data$names[data$fid %in% entity_index]), ',')

### get last and first name
entity_fst <- lapply(entity_split, '[[', 1)

### use first entity name, split by space to select those with only one entry until comma
### to get private people (usually, entities have several words)
entity_fst_split <- strsplit(as.character(entity_fst), ' ')

### get fid of entities, where we are relatively certain
entity_prp <- entity_index[lengths(entity_fst_split)>1]

### assign property code to entity_prp
data$prp[data$fid %in% entity_prp] <- 2

################################################################################
### sanity check for uncertain cases
################################################################################

### get rows where category remains undetermined
uncertain_index <- data$fid[is.na(data$prp)]

### split string by comma and fetch first entry
uncertain_split <- strsplit(as.character(data$names[data$fid %in% uncertain_index]), ',')
uncertain_fst <- lapply(uncertain_split, '[[', 1)

uncertain_entity <- uncertain_index[grep(paste(blacklist, collapse='|'), uncertain_fst)]
uncertain_people <- uncertain_index[uncertain_index %!in% uncertain_entity]

### assign property code to people and entities
data$prp[data$fid %in% uncertain_people] <- 1
data$prp[data$fid %in% uncertain_entity] <- 2

################################################################################
### create short names
################################################################################

### get first and last name
name_split <- strsplit(as.character(data$names), ', ')
```

```
 fst <- lapply(name_split, '[[', 1)
 scd <- lapply(name_split, '[[', 2)

 ### clean text
 data$short_names <- NA
 data$short_names[data$prp==1] <- as.character(paste(scd[data$prp==1], fst[data$prp==1], sep=' '))
 data$short_names[data$prp==2] <- as.character(fst[data$prp==2])

 ##############################################################################
 ### assign additional prp categories
 ##############################################################################

 data[grep(paste(private, collapse='|'), data$short_names),"prp"] <- 3
 data[grep(paste(church, collapse='|'), data$short_names),"prp"] <- 4
 data[grep(paste(unclear, collapse='|'), data$short_names),"prp"] <- 5
 data[grep(paste(public, collapse='|'), data$short_names),"prp"] <- 2

 return(data)
}
```

Further information

Literature

Bunkus, R. & I. Theesfeld (2018) Land Grabbing in Europe? Socio-Cultural Externalities of Large-Scale Land Acquisitions in East Germany. *Land,* 7. DOI: 10.3390/land7030098

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. DOI: 10.3368/le.84.4.573

Desmarais, A. A., D. Qualman, A. Magnan & N. Wiebe (2015) Land grabbing and land concentration: Mapping changing patterns of farmland ownership in three rural municipalities in Saskatchewan, Canada. *Canadian Food Studies/La Revue canadienne des études sur l'alimentation,* 2**,** 16-47. DOI: 10.15353/cfs-rcea.v2i1.52

--- (2017) Investor ownership or social investment? Changing farmland ownership in Saskatchewan, Canada. *Agriculture and Human Values,* 34**,** 149-166. DOI: 10.1007/s10460-016-9704-5

Forstner, B., A. Tietz, K. Klare, W. Kleinhanss & P. Weingarten. 2011. *Aktivitäten von nichtlandwirtschaftlichen und überregional ausgerichteten Investoren auf dem landwirtschaftlichen Bodenmarkt in Deutschland : Endbericht*. Braunschweig : vTI.

Landratsamt Märkisch-Oderland. 2019. Landkreis Märkisch-Oderland - Statistischer Jahresbericht 2018. Seelow.

Nickerson, C., M. Morehart, T. Kuethe, J. Beckman, J. Ifft & R. Williams. 2012. Trends in US farmland values and ownership. In *Economic Information Bulletin Number 92*. U.S. Dept. of Agriculture, Economic Research Service.

Tietz, A. 2019. Bodengebundene Einkommensteuern in einer strukturschwachen ländlichen Gemeinde (Supra-regionally active shareholders in East-German agricultural companies: Trends up to 2017). In *Thünen Report 69*. Braunschweig: Johann Heinrich von Thünen Institute, Federal Research Institute for Rural Areas, Forestry and Fisheries.

Trautvetter, C. & M. Henn. 2020. Keine Transparenz trotz Transparenzregister: Ein Rechercheberichт zu Anonymität im Berliner Immobilienmarkt. Berlin.

van der Ploeg, J. D., J. C. Franco & S. M. Borras (2015) Land concentration and land grabbing in Europe: a preliminary analysis. *Canadian Journal of Development Studies,* 36**,** 147-162. DOI: 10.1080/02255189.2015.1027673

Contact

Daniel Müller[1,2,3,*], Philippe Rufin[2], Marcel Schwieder[2,4]

1 Leibniz Institute of Agricultural Development in Transition Economies (IAMO), Theodor-Lieser-Straße 2, 06120 Halle (Saale), Germany

2 Geography Department, and

3 Integrative Research Institute on Transformations of Human-Environment Systems (IRI THESys), Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

4 Thünen Institute of Farm Economics, Johann Heinrich von Thünen Institute, German Federal Research Institute for Rural Areas, Forestry and Fisheries, Bundesallee 63, 38116 Braunschweig

* Corresponding author: mueller@iamo.de

Acknowledgement

We are grateful to the Federal Office for Agriculture and Food (Bundesanstalt für Landwirtschaft und Ernährung, BLE), which funded this research project (reference number 324-06.01-2818HS002) in order to provide scientific advice to the Federal Ministry of Food and Agriculture (Bundesministerium für Ernährung und Landwirtschaft, BMEL).

Berlin, May 5, 2021