A bird's eye view of farm size and biodiversity
Noack, Frederik; Larsen, Ashley; Kamp, Johannes; Levers, Christian

published in
American Journal of Agricultural Economics
2022

DOI (link to publisher)
10.1111/ajae.12274

document version
Publisher's PDF, also known as Version of record
document license
Article 25fa Dutch Copyright Act

Link to publication in VU Research Portal

citation for published version (APA)

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:
vuresearchportal.ub@vu.nl
A bird’s eye view of farm size and biodiversity: The ecological legacy of the iron curtain

Frederik Noack1 | Ashley Larsen2 | Johannes Kamp3,4 | Christian Levers5

1Food and Resource Economics, University of British Columbia, Vancouver, British Columbia, Canada
2Bren School of Environmental Science & Management, University of California, Santa Barbara, Santa Barbara, California, USA
3Department of Conservation Biology, University of Göttingen, Göttingen, Germany
4Dachverband Deutscher Avifaunisten (DDA), Münster, Germany
5Department of Environmental Geography, Institute for Environmental Studies, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Abstract
Agriculture is a major threat to global biodiversity. A common claim is that large-scale agro-industrial farming is mainly responsible for the biodiversity decline, while smaller family farms are more wildlife friendly. Here we leverage a natural experiment along the former inner German border to estimate the causal impact of farm size on biodiversity. We combine land cover data with bird diversity data to establish the mechanisms through which farm size affects bird diversity. Our main results show that the increase in farm size at the former inner German border reduces bird diversity by 15%. The results suggest further that the decline is the result of land cover simplification rather than land use intensification.

KEYWORDS
biodiversity, farm size, iron curtain

JEL CLASSIFICATION
Q15, Q57

1 INTRODUCTION

Agricultural is one of the largest threats to global biodiversity (Maxwell et al., 2016; Tilman et al., 2017). It also provides rural incomes and food for a growing global human population. Increasing agricultural production is therefore crucial for meeting future food demand and for rural development. However, increasing agricultural production may have negative consequences for biodiversity (Kehoe et al., 2017; Zabel et al., 2019). Rural incomes, agricultural productivity, and farm size are closely related (Adamopoulos & Restuccia, 2014; Eastwood et al., 2010). Increasing farm size [Correction added on 30 March 2022, after first online publication: The table captions and notes have been corrected in this version.]
provides the economic scale for the adoption of modern technologies and agricultural mechanization (Foster & Rosenzweig, 2017). Consequently, larger farms are more mechanized and use less labor and more agrochemicals per unit of land (Larsen & McComb, 2021; MacDonald et al., 2014; Muyanga & Jayne, 2019; Noack & Larsen, 2019). Although farm size generally increases with economic development (Adamopoulos & Restuccia, 2014), its impact on biodiversity is poorly understood. Testing the impact of farm size on biodiversity is complicated by the lack of exogenous variation in farm size. Here, we leverage the discontinuous change in farm size along the former inner German border to estimate the causal impact of farm size on biodiversity. To quantify the impact of farm size on biodiversity and to establish the mechanisms that relate farm size to biodiversity, we combine bird diversity data from opportunistic citizen science and a systematic bird survey with satellite data on land cover in a regression discontinuity approach.

Farm size affects biodiversity outcomes not directly but rather through input choices and the spatial organization of agriculture. For example, large farms may use different levels of agrochemicals than small farms. Increased levels of pesticides, fertilizers, and other agrochemicals can disrupt food webs and displace less competitive plant and animal species, therefore reducing biodiversity (Geiger et al., 2010; Hautier et al., 2009; Hautier et al., 2014). Large farms also change the landscape configuration through the increased scale of the operation. Large farms are associated with large fields (Lesiv et al., 2019; Samberg et al., 2016) and therefore with lower amounts of edge habitat relative to the crop area with demonstrably negative effects on biodiversity (Clough et al., 2020; Fahrig et al., 2015; Martin et al., 2019; Sirami et al., 2019). Large farms are also associated with reduced land cover and crop diversity (Batáry et al., 2017; Ricciardi et al., 2021) with potentially negative consequences for biodiversity (Strobl 2021). Such land cover simplification arises when land or crop cover is more homogeneous within than across farms and fields. In this study we seek to understand the contributions of these different mechanisms to the impact of farm size on biodiversity.

Farm size and the associated land-use changes are not random but evolve in response to geography, economic incentives, and farmer characteristics. A major challenge for estimating causal relationships between farm size and biodiversity is therefore to separate the causal effect of farm size on biodiversity from the underlying, confounding geographical and economic variables. Regression discontinuity (RDD) approaches have been used in similar settings to study the causal impact of policies on environmental outcomes (Almond et al., 2009; Burgess et al., 2018; Englander, 2019). Here, we use the discontinuous change in farm size at the former inner German border to isolate the causal impact of farm size from the continuously changing underlying geographical variables such as climate or topography.

This approach relies on the assumption that the discontinuous change in farm size is unrelated to other variables that could affect biodiversity directly. Although there are differences in culture and geography between East and West Germany (Becker et al., 2020), we argue that these differences change gradually across the former inner German border, because the exact location of the former inner German border was largely determined by the advancement of the allied forces during World War II and historic borders that were abolished more than 100 years prior to the establishment of the former inner German border (Buchholz, 1994). This argument is supported by a lack of farm size differences between East and West Germany in the 1950s, just prior to the farm collectivization in the 1960s (Koester & Brooks, 1997), and by the smooth change of climate and topographic variables across the former inner German border (see Appendix A). Although the German reunification placed Eastern and Western German farms in the same institutional setting and market environment, the differences in farm size persisted over the decades. Farms in East Germany are still on average four times larger than farms in West Germany, with little signs of convergence. Although the share of forest and crop cover changes smoothly across the former inner German border, the increased farm size has profound implications for land use. East German farms use considerably less labor per unit of land, fields are twice as large, and land cover is significantly less diverse in East Germany than in West Germany. Unlike farm size, differences in other environmental outcomes such as air pollution disappeared quickly after the reunification of Germany in 1990 (Sugiri et al., 2006). We can therefore use the discontinuous change in farm size at the former inner German
border to estimate the causal impact of farm size on biodiversity and to identify the mechanisms that underlay this relationship.

Although farm collectivization and the creation of large industrial farms were the main drivers of the current farm size differences between East and West Germany, former land use policies could have additional effects on current land use outcomes. For example, farms on both sides of the border are dominated by family owned businesses (90% West Germany, 72% East Germany), but the share of corporate farms is larger in East Germany compared to West Germany (BMEL, 2017). This difference could be a direct consequence of the farm expropriation and collectivization in East Germany instead of an indirect result of increased farm size. Further, differences in environmental policies could reinforce land use differences between East and West Germany. However, we show that farm size is the main driver of biodiversity differences in several robustness tests.

To measure biodiversity, we use data from the German Common Breeding Birds Survey (CBBS) as well as data from opportunistic citizen science (eBird). Although the CBBS data have the advantage of being a stratified random sample with a standardized sampling procedure, the density of eBird observations is several times higher than CBBS observations, allowing us to measure changes in bird diversity in higher spatial resolution. We combine these data with detailed land cover data in 20 × 20 m spatial resolution from Preidl et al. (2020) to quantify the mechanisms that relate farm size to biodiversity.

Our main results suggest a strong and negative effect of farm size on bird diversity. Using the CBBS observations, we find a decline of general bird diversity by 15% and a decline of common cropland species by 16%. We confirm these findings using our sample of eBird observations. The results are qualitatively similar although the magnitude of the effect is increased. Based on the eBird data, we find that general bird diversity declines by 20% at the former inner German border, whereas diversity of common cropland species declines by 28%.

These results suggest (1) that the impact of farm size on bird diversity is substantial, (2) that cropland species are more affected by changes in farm size than general bird diversity, and that (3) the estimates based on eBird data are larger than those based on CBBS data. These differences between the results based on the eBird and CBBS data are unlikely to stem from species selection bias in the eBird data because the difference remains when we restrict the observations to the same set of common cropland species. However, we find substantial differences in the land cover between eBird and CBBS samples suggesting spatial selection as a potential explanation for the differences in the results.

To allocate the overall effect of farm size to specific mechanisms, we add land cover controls to the RDD framework. Adding field size, land use intensity, crop diversity, and land cover diversity as controls shows that the negative impact of farm size on bird diversity is almost exclusively driven by reductions in land cover diversity. In contrast to our overall results, land cover diversity does not explain the large and negative impact of farm size on cropland species. To explore these results further, we then use Germany-wide CBBS observations in a cross-sectional regression. The cross-sectional results generally confirm the RDD results: Land cover diversity has the largest impact on bird diversity followed by cropland extent.

Our results therefore suggest that land cover diversity is an important driver of biodiversity in agricultural landscapes. The results suggest further that the impact of land cover diversity on biodiversity may even outweigh the impact of land use intensification and cropland extent on biodiversity in the agricultural landscapes of Central Europe. In these landscapes, biodiversity has adapted to

---

1Large farms in the United States are also more likely to be managed as corporate farms than smaller farms (MacDonald et al., 2014) despite the absent farm collectivization.
2Wuepper et al. (2020) show that there are no significant differences in current agri-environmental policies across the former inner German border.
3Other potential biases that affect the eBird data are endogenous observer effort and spillovers between East and West Germany. These biases are less pronounced in the CBBS data because of the high and standardized sample effort (see Section 2) and the focus on breeding birds, reducing the impact of spillovers from, for example, migratory birds.
agriculture during the long cultivation history, while species that depend on large undisturbed habitats have mostly disappeared. Biodiversity in these landscapes therefore depends crucially on agricultural practices.

Agriculture is now the dominant form of land use in many parts of the world. Any changes in agricultural practices therefore have important implications for global biodiversity conservation. Conserving biodiversity within the agricultural landscape is not only important for achieving global conservation targets but also because of the ecosystem services it provides for agriculture. Local biodiversity enhances and stabilizes agricultural production through ecosystem services including pollination and natural pest control (Binder et al., 2018; Dainese et al., 2019; Garibaldi et al., 2016; Larsen & Noack, 2017; Martin et al., 2019; Noack et al., 2019). Despite these benefits, biodiversity is often neglected in the farmers’ decision. Governments therefore implement expensive agri-environmental policies to mitigate the negative impact of agriculture on the environment. The European Union spends several billion Euro every year on agri-environmental policies (Arata & Sckokai, 2016; Wätzold et al., 2016), whereas the Endangered Species Act reduces farmers' incomes and land values across the United States (Melstrom, 2019). Despite these efforts, the success of these policies is limited (Wätzold et al., 2016). Targeted agri-environmental policies can reduce conservation costs while improving biodiversity outcomes. Our results show that the negative impact of increased farm size can be mitigated by conserving land cover diversity within the agricultural landscape. In practice, this could mean incentivizing riparian buffers strips, forest patches, hedgerows, or agroforestry. Our results therefore provide new opportunities for targeted interventions to harmonize agricultural intensification with biodiversity conservation targets.

Our study contributes to two strands of literature. The first strand concerns the protection of biodiversity on private land. Although much of the economics literature focuses on the Endangered Species Act and other conservation policies (Chabé-Ferret & Subervie, 2013; Langpap & Junfje, 2017; Melstrom, 2019; Wätzold et al., 2016), our study contributes to this literature by addressing the question about the consequences of agricultural development for biodiversity conservation. A large literature in ecology and environmental science relates agricultural intensification to biodiversity outcomes, but those studies are either small-scale experiments or large-scale correlations. Two recent studies are closely related to our paper. Li et al. (2020) find a negative effect of neonicotinoid insecticides on birds diversity in the United States using an instrumental variable approach. A second study by Strobl (2021) establishes a positive correlation between crop diversity and bird diversity in the United States. However, to our knowledge, our paper is the first study to estimate the causal impact of farm size on biodiversity at the landscape scale.

The second strand of related literature estimates the impact of farm size on agricultural productivity (see Barrett et al. 2010 and Gollin, 2019 for summaries of the recent literature). This literature largely neglects the environmental consequences of increased farm size. As far as we know, only two papers at very different scales relate farm size to environmental outcomes. Using a regression discontinuity approach, Wuepper et al. (2020) find that small farms lead to more spatial heterogeneity in production and more bare soils. Their study covers a similar area as the current study but focuses on management practices as proxies for environmental outcomes. In contrast to this large-scale study, Batáry et al. (2017) use field level data to compare beetles, spiders, and plants between small and large fields. Our study bridges these studies by combining land use data from satellite images with biodiversity data from field surveys.

In the next two sections we introduce our data (Section 2) and illustrate the impact of farm collectivization on farm size and land use (Section 3). In the main section of our paper we then discuss the regression discontinuity estimation strategy and results on the impact of farm size on

---

5See Ando and Langpap (2018) for a recent summary of the economics literature.
6See Tschamktie et al. (2012) and Kremen and Merenlender (2018) for summaries of the ecology and environmental science literature.

---

On page 2 they write “Optimally, we would be able to directly measure environmental outcomes, e.g. changes in biodiversity or soil and water quality (perhaps even at the landscape level). Because such data are not available, we must rely on management practices that have environmental implications.”
biodiversity (Section 4). We support the mechanisms suggested by these results with a cross-sectional analysis of land use to biodiversity (Section 5). Section 6 concludes.

2 | DATA

To estimate the impact of farm size on biodiversity and to relate our results to potential mechanisms, we use data from a range of sources. We use bird diversity data from citizen science (eBird) (Sullivan et al., 2009) and structured bird surveys (CBBS) (Kamp et al., 2020). We then match bird sampling locations with land cover data from Preidl et al. (2020), vegetation productivity from Radeloff et al. (2019), climatic data from Karger et al. (2017), and elevations data from Danielson and Gesch (2011). CBBS bird diversity is recorded in 1 × 1 km sample areas and we use the same area around the centroid of the eBird observation to relate bird diversity to the land cover and other geographic variables. Our main analysis focuses on the former inner German border. We therefore summarize our data in 50 km bands on either side of the former inner German border. This chapter introduces the data sources in more detail and describes the processing algorithms.

2.1 | Bird diversity

2.1.1 | eBird

Being one of the world’s largest biodiversity-related citizen science projects, the eBird database contains information on bird distribution, bird abundance, and habitat use (Sullivan et al., 2009). The database is populated by volunteers who enter the time, duration, location, and mode of their bird watching activity, as well as birds seen and heard during the outing through a standardized checklist. In the following, we call all species that were recorded during one outing (a unique observer, location, and time combination), a bird diversity observation. We use the eBird Basic Dataset (EBD) for Germany, version September 2019, from https://ebird.org eBird. We use the auk package (Strimasc-Mackey et al., 2018) in R to preprocess the data, following the best practices examples for eBird data (Johnston et al., 2020) regarding data cleaning and filtering. To reduce the variation in observation effort between observations, we restrict our sample to observations with a duration of less than 5 h, a distance traveled during the observation of less than 5 km, and with 10 or fewer observers. Further, to control for observer experience and effort, we include experience measured by the number of observations entered previously by the same observer into the eBird database and observer effort measured by the duration of the observation as controls in all regression specifications. Finally, we only include complete observations, observations with some cropland (>0) in the 1 × 1 km sample areas, and observations between 2015 and 2018 to match them with the land cover data. After this data cleaning process, 2371 bird diversity observations remain in our sample within the 50 km band on either side of the former inner German border. Panel B in Figure 3 shows the distribution of the bird diversity observations within the 50 km band on both sides of the former inner German border.

2.1.2 | Matching

The eBird observations in East and West Germany differ in terms of their timing, their observer characteristics and their observation effort. We therefore preprocess the eBird data to make the observations in East and West Germany comparable in terms of their timing, their observer

We also include the 2018 observations in order to increase the sample size because eBird use in Germany grew substantially over this time period.
characteristics, their observation effort, and the distance to the former inner German border, following the method suggested by Ho et al. (2007). We use nearest neighbor matching with logistic distance as suggested by Ho et al. (2011). We restrict the sample to observations within 50 km distance to the former inner German border. The matching is based on the distance of the observation to the former inner German border, the duration of the observation, the experience of the observer as well as the year and the month of the observation. We include duration and experience in logs to account for a declining marginal impact of experience and effort on outcome. We show results based on the non-matched data in a robustness test. After matching, 579 observations remain.

2.1.3 | German common breeding bird survey (CBBS)

In the German Common Breeding Bird Survey, volunteers map breeding birds within quadratic 1 × 1 km sample areas (Kamp et al., 2020). The sample areas are surveyed four times in the period from March 10 to June 20 to ensure that all breeding birds are recorded. Repeated sightings of the same species within the same location over the four survey rounds within one breeding season are combined into territories at the end of the season using the territory mapping method. For many species, territories are equivalent to breeding pairs although there are exceptions such as polygamous species where a territory consist of a male and several females. For the CBBS data, one bird diversity observation is defined as the number of breeding bird territories (pairs) within the 1 × 1 km sample area during one breeding season (a unique sample area–year combination). More than 2600 quadratic sample areas of 100 ha (1 × 1 km) are distributed across Germany. When the monitoring scheme was set up in 2005, plots were selected randomly stratified, based on a two-level stratification. The first stratum are Germany’s “environmental regions,” that is, spatial units of similar environmental characteristics such as soil, climate, and vegetation (Kamp et al., 2020). The second stratum is land cover with the classes arable land, rare crops (e.g., grapes and vines), grassland, forest, settlements, and a combined class of rarer land-cover types (Kamp et al., 2020). Most sample areas were surveyed by the same observer for several years. We include only data between 2015 and 2017 to match them with the land cover data and restrict the sample to areas with at least some (> 0) crop cover within the 1 × 1 km sample areas. After this process we remain with 2723 bird diversity observations (sample area–year combinations) for Germany of which 561 fall within the 50 km band on both sides of the former inner German border. Panel B in Figure 3 shows the distribution of the bird diversity observations within the 50 km band on both sides of the former inner German border.

2.1.4 | Diversity measure

Our main measure of bird diversity is species richness, but we also report results based on the inverse Simpsons Diversity Index (SDI) in a robustness test. This specification of the SDI is also called effective number of species (Hill, 1973) because it equals species richness (the number of species) when all species are equally abundant, and it declines with increasing inequality between species’ abundances. Our rationale for choosing bird richness as our main measure is to make the results between CBBS and eBird data directly comparable. CBBS reports abundance of all species, whereas abundance data are often absent in the eBird data.

8Most biodiversity measures combine the number of species and their abundance. The most commonly used diversity measures (SDI, Shannon index and species richness) can be defined by

\[ D = \left( \sum_{i=1}^{n} a_i^\alpha \right)^{-1/\alpha} \]

where \( \alpha \geq 0 \), \( n \) is the number of species, and \( a_i \) is the relative abundance of species \( i \) (Hill, 1973). The main difference between the specific indices is the weight placed on abundance compared to mere presence-absence information. For \( \alpha = 0 \) the index equals species richness, that is, the number of unique species in a sample area, whereas for \( \alpha = 2 \) the index equals the SDI.
Figure 1 summarizes all bird diversity observations of our sample within the 50 km band on both sides of the former inner German border. Average species richness is higher in the CBBS data (Figure 1 panel B) than in the eBird data (Figure 1 panel A). Although the observation effort is standardized in the CBBS data, it is not in the eBird data.

2.1.5 | Cropland species

We define a set of cropland species using the complete CBBS data set for Germany for the years 2015 to 2017. In a first step we estimate the relationship between crop cover in the sample areas and the number of breeding pairs for all species with more than 250 observations in individual regressions:

\[
\text{territories}_{ijst} = \alpha_i \text{crop}_{js} + \theta_t + \gamma_s + \epsilon_{ijst}
\]

where “territories” is the number of territories of species \(i\) in sample area \(j\) in state (Bundesland) \(s\) in year \(t\), crop is the crop cover in percent in sample area \(j\) in the year 2016 and the remaining variables are state and year fixed effects and an error term. We therefore only use the variation of crop cover and bird abundance within states to estimate the relation between crop cover and species abundance. We cluster standard errors at the site level to account for serial correlations. In a second step, we define species as cropland species if \(\alpha_i > 0\) and a \(p\)-value below 0.01, that is, a statistically significant positive relationship between species abundance and crop cover within the sample area. The list of the species is given in Appendix B.

2.2 | Land cover

To match land cover to bird diversity data, we summarized land cover categories within our 1 \(\times\) 1 km sample areas for both eBird and CBBS observations. The 1 \(\times\) 1 km area represents the immediate habitat of the sighting. In a robustness test we use 6 \(\times\) 6 km squares to represent the wider habitat of the sighting (Mitchell et al., 2001; Semper-Pascual et al., 2018).

We use the most recent and accurate agricultural cropland cover classification for Germany available (Preidl et al., 2020) to measure land cover. The data layer was generated for the year 2016 using Sentinel-2A satellite imagery (20 \(\times\) 20 m spatial resolution) and a random forest classifier.
It achieved an overall accuracy of 88% for the following 23 land-cover classes: winter wheat, spelt, winter rye, winter barley, spring wheat, spring barley, spring oat, maize, legumes, rapeseed, leeks, potatoes, sugar beets, strawberries, stone fruits, vines, hops, asparagus, grassland, buildings, water bodies, other vegetation, and forest. This accuracy does not differ between East and West Germany (see table 3 in Preidl et al. [2020]).

### 2.2.1 | Land cover and crop diversity

To measure land cover diversity, we aggregate the original land cover categories to the categories: grassland, cropland, forest, buildings, water, and other land cover. We use the SDI to measure land cover diversity because it takes the relative abundance of land cover types into account, which is important to characterize bird habitat. Similar to land cover, we also use the SDI to measure crop diversity but using the disaggregated crop cover categories.

### 2.2.2 | Field size

We approximate field size with the size of crop patches due to the lack of peer-reviewed, spatially explicit, and high-resolution field size data for Germany, following the argumentation by Weissteiner et al. (2016). Crop patches are contiguous, homogenous crop cover units. Crop cover includes the following categories: winter wheat, spelt, winter rye, winter barley, spring wheat, spring barley, spring oat, maize, legumes, rapeseed, leeks, potatoes, sugar beets, strawberries, stone fruits, vines, hops, and asparagus.

To calculate field size for the complete 50 km band east and west of the former inner German border, we first convert the categorical land-cover map (Preidl et al., 2020) into binary maps for each crop type class. Second, we convert the resulting raster maps into polygon shapefiles, with separate polygons representing contiguous patches for each crop type class. Third, we calculate the area of each polygon. Last, we calculate the distance of the centroid of each polygon to the former inner German border and merge all crop type classes into one data table. To exclude solitary pixels as well as very small or large fields, we trim our data to crop fields larger than 0.36 ha (i.e., nine pixels in any spatially contiguous shape) and smaller than 1000 ha (i.e., 25,000 pixels in any spatially contiguous shape). We further removed all fields located closer than 1 km from the former inner German border to remove potential trans-border fields, resulting in a final sample size of 666,945 fields.

### 2.2.3 | Summary land cover

Figure 2 summarizes the broad land cover categories within the 1 × 1 km CBBS sample areas and the 1 × 1 km area around the centroid of each eBird observation. The figure shows that the eBird and CBBS data differ with respect to their surrounding land cover. First, more eBird observations are located in residential areas dominated by buildings (Figure 2, panel A) compared to the CBBS observations (Figure 2, panel B). Second, the relative difference between the land cover in East and West Germany is larger in the eBird observations compared to the CBBS observations except for the otherland cover category.

### 2.3 | Additional data

#### 2.3.1 | Farm size, yields, labor and ownership

We use data on crop yields, agricultural labor, farm ownership, and farm size on district (Kreis) level from the German harvest statistics (Emtestatistik) and the agricultural census (Allgemeine
**FIGURE 2** Distribution of land cover

Notes: (a) eBird sample, (b) CBBS sample.

**FIGURE 3** Farm size, bird diversity observations and field size

Notes: Panel A shows district level mean farm size in Germany for the year 2016. The dark lines depict the former inner German border and the 50 km bands on both sides of the former inner German border. Panel B shows the distribution of bird diversity observations within the 50 km bands on both sides of the former inner German border. Panel C shows field size (patch size) for all crop fields within the 50 km bands on both sides of the former inner German border.
Agricultural inputs and outcomes across the former inner German border in Section 3 and for robustness tests. The agricultural census data are for 2016 except for the labor data, which are only available for 2010.

2.3.2 | Enhanced vegetation index (EVI)

We use the EVI as a proxy for land-use intensity due to its correlation with crop yields (Bolton & Friedl, 2013; Burke & Lobell, 2017) in combination with land cover shares as controls in our regression specification. EVI is more sensitive to changes in areas of high biomass compared to NDVI. Here, we use the cumulative EVI from MODIS 1000-m vegetation productivity data sets for 2003–2014, generated as a Dynamic Habitat Index (Radeloff et al., 2019). We match the EVI data to the bird data using the nearest EVI pixel to the bird sample area centroid.

2.3.3 | Climate and topography

We use CHELSA bioclimatic variables (Karger et al., 2017) to represent climatic conditions for CBBS and eBird observations. We use mean annual temperature (Bio1) and annual precipitation (Bio12), which influence bird diversity patterns as they determine water availability and energy available for plant growth (Elsen et al., 2020; Karger et al., 2017). We further use the GTOPO30 digital elevation model at 1 × 1 km spatial resolution (Danielson & Gesch, 2011) to control for the influence of topography on bird diversity. We match these data to the bird diversity observations using the minimum distance to the sample area centroid.

2.3.4 | Protected areas

In one robustness test we control for the distance from the sample area centroid to the nearest protected area. We use the protected area shapefiles from the European Union’s Common Database on Designated Areas.

2.4 | Data summary

Figure 3 illustrates our data. Panel A shows the location of the former inner German border and the 50 km buffer that defines our study area. The farm size difference between East and West Germany becomes clearly visible in the shading of the districts. Panel B shows that the eBird and CBBS bird diversity observations are distributed throughout the whole study area. Panel C shows the distribution of cropland and field sizes. Cropland is equally distributed across the whole study area except for a few mountain ranges, lakes, and larger cities (dark areas). Fields in West Germany are generally smaller than fields in East Germany.

3 | THE LAND USE LEGACY OF FARM COLLECTIVIZATION

We find a large and discontinuous increase in farm and field size at the former inner German border as a legacy of agricultural collectivization in East Germany after the Second World War. Farm size in Germany also varies because of other factors, including topography and inheritance pattern, but this source of variation is small compared to the variation of farm size due to farm collectivization. Although the terrain in Southern Germany is largely mountainous or hilly, it is largely flat in
Northern Germany. These topographic differences led to smaller farms in Southern Germany and larger farms in the Northern Germany. Differences in land inheritance pattern between Northern and Southern Germany magnified these differences. Although the agricultural land in Southern Germany was divided equally among the heirs (“Realeitung”) leading to smaller farms and fields, the farmland in Northern Germany was mainly bequeathed to the oldest heir (“Anerbenrecht”), leaving the farm structure intact. These differences in topography and inheritance pattern led to considerable farm size differences between Southern and Northern Germany. According to the German agricultural census of 2016, average farm size in northern West Germany was 71 hectares, whereas it was only 35 hectares in southern West Germany. However, these differences are small compared to the differences between East and West Germany that resulted largely from farm collectivization in the formally socialist East Germany. According to the 2016 German Agricultural Census, average farm size in West Germany was 45 hectares, whereas it was 223 hectares in East Germany (see Panel A in Figure 3).

These large farm size differences between East and West are not the result of differences in topography or inheritance pattern but largely due to land collectivization. Prior to farm collectivization, farm size was similar between East and West Germany. In 1950, the average farm size was 8.2 hectares in West Germany and 8.8 hectares in East Germany (Koester & Brooks, 1997). After 1950, the socialist government in East Germany implemented agrarian reforms that led to large-scale and rapid land expropriation and collectivization. In 1950 almost all farms were private, but only 10 years later, in 1960, more than 90% of the East German farms were collectivized (Koester & Brooks, 1997). During this process, private farms were aggregated to large collective farms. After the reunification of Germany in 1990, farms in East Germany were privatized and since then have operated in a very similar political and economic environment as farms in West Germany. The process of privatization in East Germany was relatively quick. Following this, the initial transformation of large collective farms to private enterprises, farm size distribution has remained stable over the past two decades (Figure 4), with large farms in East Germany and relatively small farms in West Germany (Figure 4).

This change in farm size at the former German border provides a natural experiment to test the impact of farm size on biodiversity. However, the approach relies on the assumption that the location of the former inner German border is not correlated with other variables that could also affect biodiversity. We present regression discontinuity plots for geographic variables including climate, elevation, and soil characteristics in Appendix A. These geographic variables exhibit no discontinuous change at the former inner German border. Only precipitation declines continuously from West to East but shows no abrupt change at the border. We therefore include precipitation as control in all regression specifications.

Farm size has no direct impact on bird diversity. Instead, it affects bird diversity through a range of mechanisms including agricultural mechanization, different levels of input use, field size, and land cover simplification (including crop diversity). The relationship between farm size and these mechanisms has been documented for a range of countries (see e.g. Ricciardi et al. 2021, Levin 2006). In the following we discuss their differences across the former inner German border.

Figure 5 illustrates the field size differences between East and West Germany based on all fields within a 50 km band on both sides of the former inner German border (666,945 observations). It shows that fields in East Germany are, on average, almost twice as large as in West Germany. Despite the large differences in farm and field size, we find no discontinuous change in land use intensity at the former inner German border when using the EVI index as proxy for land use intensity (Appendix C). This finding is also supported by Panel A of Figure 6, which plots district level wheat yields (the most common crop in the study area) against the districts’ distances to the former inner German border. The figure suggests that wheat yields are generally higher in West Germany compared to East Germany, although the differences are small. In contrast to wheat yields, labor inputs between East and West Germany differ largely. Panel B of Figure 6 suggest that West German farms use on average twice as much labor per unit of land as East German farms. Figure 6 therefore suggests that potential efficiency gains from increased farm size are likely to result from labor saving technologies rather than increased yields.

Another land use implication of increased farm and field sizes is the loss of land cover diversity (see e.g. Ricciardi et al. 2021). Fields and farms are typically managed as relatively homogeneous
Dividing the same amount of land among fewer farms and fields therefore reduces land cover diversity. Figure 7 illustrates the relationship between crop diversity and the distance to the former inner German border for all 1 km grid cells with crop cover within the 50 km bands on both sides of the border.

**Figure 4** Farm size in east and West Germany

Notes: Panel A plots the mean district level farms size against the distance of the district centroid to the former inner German border. The size of the points is relative to the district’s agricultural area. The lines are the mean farm size in East and West Germany (agricultural area weighted district means). Panel B shows the district level farm size distribution in East and West Germany over time.

**Figure 5** Field size and distance to the former inner German border

Notes: The dots are sample means in 1 km distance bins from the former inner German border. The bars are 95% confidence intervals. West Germany is to the left of the former inner German border, East Germany is to the right of the former inner German border. The sample consists of all crop patches larger than 3 pixels (0.36 ha) in a 50 km distance band on both sides of the former inner German border.
The main aim of this paper is to estimate the causal impact of farm size on bird diversity and to identify plausible mechanisms for this relationship. To do so, we exploit the sharp farm size discontinuity at the former inner German border in a regression discontinuity design. Figure 8 visualizes the distribution of bird diversity across the former inner German border using the preprocessed eBird data. The figure suggests a discontinuous change in bird diversity at the former inner German border as well as a gradual increase of bird diversity from West to East.

4 | FARM SIZE AND BIRD DIVERSITY

The main aim of this paper is to estimate the causal impact of farm size on bird diversity and to identify plausible mechanisms for this relationship. To do so, we exploit the sharp farm size discontinuity at the former inner German border in a regression discontinuity design. Figure 8 visualizes the distribution of bird diversity across the former inner German border using the preprocessed eBird data. The figure suggests a discontinuous change in bird diversity at the former inner German border as well as a gradual increase of bird diversity from West to East.

4.1 | Estimation

In this section, we formalize the regression discontinuity approach illustrated in Figure 8. We also develop the approach to identify the mechanisms that relate farm size to bird diversity.

Following Gelman and Imbens (2019), we use local linear regressions with a triangular kernel and optimal bandwidth selection based on the mean squared error criterion. The baseline regression discontinuity equation is

\[ y_{jt} = f(\text{running}_{j}) + \beta_{\text{East}_j} + Z_j \eta + \varepsilon_{jt} \]  

(1)
**FIGURE 7**  Crop diversity and distance to the former inner German border  
*Notes:* Crop diversity is measured in all $1 \times 1$ km grid cells with crop cover using the SDI. The dots are sample means in 1 km distance bins from the former inner German border. The bars are 95% confidence intervals. The red line shows mean field size for East and West Germany separately. West Germany is to the left of the former inner German border, East Germany is to the right of the former inner German border. The sample consists of all crop patches larger than 3 pixels (0.36 ha) in a 50 km distance band on both sides of the former inner German border.

**FIGURE 8**  Bird species richness and distance to the former inner German border  
*Notes:* Regression discontinuity plot of residualized bird diversity using eBird data between 2015 and 2018 in a 50 km band on both sides of the former inner German border. The dots represent mean residualized bird diversity in 5 km distance bins. To residualize bird diversity, we run a regression of bird diversity on observer effort, observer experience, border, precipitation, year and quarter (see baseline specification in the next section). The figure uses the residuals of this regression. The gray line is a local linear regression with triangular kernel. West Germany is to the left, East Germany is to the right of the former inner German border.
where $y_{jt}$ is the outcome in sample area $j$ in year $t$, $(frunning)_j$ is a function of the running variable (local linear), $East_j$ is a dummy for East Germany, $Z_j$ are controls, and $\epsilon_{jt}$ is the error term. In our baseline specification for the eBird data, the controls include a dummy for the former inner German border (a band of 2.5 km on both sides of the former inner German border), average precipitation levels (see Appendix A), the log observer effort and experience, as well as year and quarter fixed effects. We include a border dummy to take the low land-use intensity at the former inner German border as a result of former border security restrictions and conservation measures after the German reunification into account. We include the experience and observation effort in logs to account for concavities in learning and species area relationships.

We subsequently add additional controls to establish the mechanism through which farm size affects bird diversity. This approach builds on the assumption that the estimated discontinuity of bird diversity would converge to zero if the added controls explained the outcome. For example, if the discontinuous change in field size at the former inner German border (Figure 5) explained the decline in bird diversity, then adding field size as a control would explain the changes in bird diversity, and the former inner German border itself had no further impact on the outcome. These controls include land cover types, land cover diversity, field size, crop diversity, and EVI as a measure of land use intensity. We measure these land cover variables in the $1 \times 1$ km sample areas in our main specification and in a $6 \times 6$ km buffer around the centroid of the observation in a robustness test. We add these controls successively to determine their individual contribution to the biodiversity response to farm size.

In an additional robustness test we implement an alternative approach to estimating the causal mechanisms that relate farm size to bird diversity. We follow Acharya et al. (2016) by estimating the controlled direct effect, that is, the effect of farm size on bird diversity after setting the mediators (field size, land cover diversity, etc.) equal to zero. We estimate the demediation function (the function used to set the mediators equal to zero) within the respective bandwidth using an ordinary least squares regression.

Former land use differences between East and West Germany may have affected current land use through channels other than farm size. For example, the share of natural vegetation and subsequently the number of protected areas may differ between East and West Germany. Although we control for the vegetation type directly in all but the baseline specifications, we additionally control for the distance to the nearest protected area in an additional robustness test (Appendix G).

To select our optimal bandwidth, we use mean squared error bandwidth selection procedure, following the suggestion by Imbens and Kalyanaraman (2012) with robust bias corrected confidence intervals as suggested by Calonico et al. (2014, 2020), and cluster our standard errors on district level. In two related robustness tests, we double the bandwidth and use a bandwidth based on coverage error probability bandwidth selection (Calonico et al., 2020).

### 4.2 Results

In the following we present our general results based on the eBird data and the CBBS data. We then present results on the diversity of cropland bird.

#### 4.2.1 eBird

Using data from eBird, we find a large and negative impact of farm size on bird species richness (Table 1). We use the optimal bandwidth from the baseline specification for all specifications of Table 1 to make the estimates directly comparable. Our baseline specification indicates that bird species richness declines by 8.5 species at the former inner German border. This is equivalent to a 20%

---

$^9$Forests are the natural vegetation in our study area (Bohn & Gollub, 2006), which is included as one of the land cover categories in the second regression specification.
decline at the former inner German border compared to the predicted bird diversity at the western side of the former inner German border. Including land cover, field size, EVI, and crop diversity as controls reduces the estimate to 6.4 species or 14%. However, including land-cover diversity as control renders the estimate statistically indistinguishable from zero. Our measure of land cover characteristics describe the bird habitat of the observation only imprecisely because the eBird observations are not restricted to a standardized area. Using land cover characteristics within a 6 km sampling square around the observation centroid does not change our findings qualitatively, however (Appendix E).

In further robustness checks, we use the non-matched data (Appendix F), control for the distance to the nearest protected area (Appendix G), use different bandwidths (Appendix D), add additional observer controls (Appendix H), and use the method suggested by Acharya et al. (2016) to estimate the impact of the mechanisms (Appendix I). Although the results remain qualitatively unchanged in these robustness tests, a placebo test indicates that the results based on the eBird data may be unreliable (Appendix J).

Another major concern with citizen science data is the selection of study sites, which is strongly affected by landscape composition. In contrast, the CBBS is based on a stratified random sample and includes the exact outline of the sampling area. Although the number of samples in proximity to the former inner German border is lower, CBBS allows us to exactly match the bird observation to land cover characteristics. Consequently, the CBBS data provide a means to evaluate the robustness and validity of the results based on the eBird data.

### 4.2.2 | CBBS

Similar to the results based on the eBird data, we find a negative and significant impact of farm size on species richness using the CBBS data. We follow the same estimation strategy as for the eBird data but drop the observer effort and experience as well as season dummies because the observer effort and timing is identical across all CBBS observations.

Our baseline result suggests that bird species richness declines by 5.76 species or 16% at the former inner German border (Table 2). Including land cover, field size, crop diversity, and EVI successively as controls has no major impact on the magnitude of the estimate. In contrast, including land cover diversity as a control reduces the estimate significantly. We interpret these results as further evidence that changes in overall land cover complexity explain the reduction in bird species richness in response to increased farm size.

### TABLE 1 Farm size and bird diversity (eBird)

<table>
<thead>
<tr>
<th>Specification</th>
<th>Estimate</th>
<th>SE</th>
<th>p</th>
<th>BW</th>
<th>NW</th>
<th>NE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>-8.5</td>
<td>1.7</td>
<td>0.000</td>
<td>7</td>
<td>41</td>
<td>77</td>
</tr>
<tr>
<td>Land cover</td>
<td>-8.7</td>
<td>2.3</td>
<td>0.000</td>
<td>7</td>
<td>41</td>
<td>77</td>
</tr>
<tr>
<td>Field size</td>
<td>-8.2</td>
<td>2.4</td>
<td>0.000</td>
<td>7</td>
<td>41</td>
<td>77</td>
</tr>
<tr>
<td>Crop diversity</td>
<td>-7.6</td>
<td>2.0</td>
<td>0.000</td>
<td>7</td>
<td>41</td>
<td>77</td>
</tr>
<tr>
<td>EVI</td>
<td>-6.4</td>
<td>2.0</td>
<td>0.002</td>
<td>7</td>
<td>41</td>
<td>77</td>
</tr>
<tr>
<td>Land cover diversity</td>
<td>-1.5</td>
<td>1.9</td>
<td>0.4</td>
<td>7</td>
<td>41</td>
<td>77</td>
</tr>
</tbody>
</table>

Note: Regression discontinuity results based on eBird data. The columns report the coefficient estimates, the standard errors (SE), the p-value (p), the bandwidth in km (BW) as well as the sample size in West Germany (NW) and in East Germany (NE). The running variable (distance from the border) is negative for West Germany and positive for East Germany. The baseline specification includes a border dummy, annual precipitation levels, log(experience), log(observation duration), year fixed effects and quarter fixed effects as controls. We add additional controls within a 1 km area around the sample centroid as control in the subsequent specifications. All controls are additive. ‘Land cover’ adds the percentage of land cover categories (crop, grassland, forest, water, built up), ‘Field size’ adds additionally the mean field size within the sample area as control, ‘Crop diversity’ adds additionally crop diversity as control, ‘EVI’ adds the EVI index as control while ‘Land cover diversity’ adds land cover diversity as control. Standard errors are clustered at the district level. The optimal bandwidth is based on the baseline specification.
Similar to the results based on the eBird data, we vary the bandwidth (Appendix D), add the distance to protected areas as controls (Appendix G), and use the method suggested by Acharya et al. (2016) to estimate causal mechanisms (Appendix I) as robustness tests. A placebo test further reveals that the estimates are relatively stable and converge toward zero as the RDD cutoff is moved further away from the former inner German border (Appendix J).

To test for the impact of farm size and farm management directly, we add district level farm size, the share of family owned farms, and agricultural labor as controls to our baseline regression specification (Appendix K). Adding farm size as a control reduces the direct impact of the former inner German border on bird diversity to zero, thus supporting our interpretation of the results.

The relative magnitude of the estimated impact of farm size on bird diversity is larger based on the eBird data than based on the CBBS data. This difference may be explained by selection bias in the eBird data either with respect to sites or species, or by endogenous observation effort or spillovers. For example, observation effort in the eBird data may respond to farm size or bird diversity. This would lead to an overestimation of the negative impact of farm size on biodiversity if the observation effort declines with reduced bird diversity or increasing farm size. Including additional observer effort controls in the eBird RDD specifications leads to a convergence of the results (Appendix H). In addition, birds are mobile and may be recorded in or over unsuitable habitat during migration. This would lead to an underestimation of the impact of farm size on bird diversity. However, these biases are less pronounced for the CBBS data because of the high and standardized observation effort (see Section 2) and the focus on birds that actively breed in the sample areas. We explore potential species and site selection biases further in the next subsection.

### 4.2.3 | Cropland birds

To limit the influence of selection bias stemming from the recording (or not recording) of certain species in the eBird data, we estimate the impact of farm size on species richness using the same estimation strategy as before but restricting our sample to the same set of cropland bird species in both data sets (see Section 2). Overall, we find a negative and significant decline in the number of cropland bird species in response to increased farm size at the former inner German border based both data sets (Tables 3 and 4). We focus our discussion on the land cover specification because of the higher precision of the estimates compared to the baseline specification, especially for the CBBS data.
The results based on the eBird data (Table 3) suggest a decline of one cropland bird species at the former inner German border in the land cover specification. The decline of one cropland bird species implies a bird diversity decline of 28% relative to the predicted farmland diversity at the western side of the former inner German border.

Table 4 reports the results for cropland birds based on the CBBS data. The estimates are very similar in absolute terms to the results based on the eBird data, but the precision of the estimates is low in the baseline specification. Adding land cover as control increases the precision of the estimate. The reduction of cropland birds by two at the former inner German border in the land cover specification implies a reduction of cropland bird species by 16% compared to bird species diversity at the western side of the former inner German border.

These results suggest that farm size affects cropland bird diversity more than general bird diversity. However, the results also suggest that species selection does not drive the difference between results based on CBBS and eBird data. In contrast to the previous subsections, land cover diversity does not explain the impact of farm size on bird diversity. Instead, our results suggest that farm size affect the diversity of cropland birds through channels other than our controls.

### Table 3  Farm size and farmland birds (eBird)

<table>
<thead>
<tr>
<th>Specification</th>
<th>Estimate</th>
<th>SE</th>
<th>p</th>
<th>BW</th>
<th>NW</th>
<th>NE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>−0.92</td>
<td>0.27</td>
<td>0.001</td>
<td>8</td>
<td>43</td>
<td>78</td>
</tr>
<tr>
<td>Land cover</td>
<td>−1.02</td>
<td>0.25</td>
<td>0.0001</td>
<td>8</td>
<td>43</td>
<td>78</td>
</tr>
<tr>
<td>Field size</td>
<td>−1.02</td>
<td>0.25</td>
<td>0.0001</td>
<td>8</td>
<td>43</td>
<td>78</td>
</tr>
<tr>
<td>Crop diversity</td>
<td>−1.02</td>
<td>0.25</td>
<td>0.0000</td>
<td>8</td>
<td>43</td>
<td>78</td>
</tr>
<tr>
<td>Intensity</td>
<td>−0.77</td>
<td>0.26</td>
<td>0.003</td>
<td>8</td>
<td>43</td>
<td>78</td>
</tr>
<tr>
<td>Landscape</td>
<td>−0.62</td>
<td>0.26</td>
<td>0.02</td>
<td>8</td>
<td>43</td>
<td>78</td>
</tr>
</tbody>
</table>

Note: The impact of farm size on farmland bird species richness using eBird data. The columns report the point estimate for the differences in bird diversity between East and West Germany (Est) at the border, the standard errors (SE), the p-value (p), the bandwidth in km (BW) and the sample size in West Germany (NW) and in East Germany (NE). The running variable is negative for West Germany and positive for East Germany. The baseline specification includes a 5 km border dummy as well as log(experience) and log(duration) of the observation as controls. The subsequent specification add additional controls. ‘Land cover’ adds the percentage of land cover categories (crop, grassland, forest, water, built up) within a 1 x 1 km square around the sample centroid as control, ‘Field size’ adds additionally the field size as control, ‘Crop diversity’ adds additionally crop diversity as control, ‘EVI’ adds the EVI index as control while ‘Land cover diversity’ includes additionally general land cover patch size and land cover diversity as controls. Standard errors are clustered at the district level.

### Table 4  Farm size and farmland birds (CBBS)

<table>
<thead>
<tr>
<th>Specification</th>
<th>Estimate</th>
<th>SE</th>
<th>p</th>
<th>BW</th>
<th>NW</th>
<th>NE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>1.3</td>
<td>1.6</td>
<td>0.4</td>
<td>19</td>
<td>74</td>
<td>138</td>
</tr>
<tr>
<td>Land cover</td>
<td>−2.0</td>
<td>0.7</td>
<td>0.002</td>
<td>19</td>
<td>86</td>
<td>150</td>
</tr>
<tr>
<td>Field size</td>
<td>−2.3</td>
<td>0.6</td>
<td>0.000</td>
<td>19</td>
<td>86</td>
<td>150</td>
</tr>
<tr>
<td>Crop diversity</td>
<td>−2.2</td>
<td>0.6</td>
<td>0.000</td>
<td>19</td>
<td>86</td>
<td>150</td>
</tr>
<tr>
<td>Intensity</td>
<td>−2.2</td>
<td>0.6</td>
<td>0.000</td>
<td>19</td>
<td>86</td>
<td>150</td>
</tr>
<tr>
<td>Landscape</td>
<td>−1.9</td>
<td>0.6</td>
<td>0.003</td>
<td>19</td>
<td>86</td>
<td>150</td>
</tr>
</tbody>
</table>

Note: The impact of farm size on farmland birds using CBBS data for the years from 2015 to 2017. The columns report the estimated differences in bird diversity between East and West Germany, the standard errors (SE), the p-value (p), the bandwidth in km (BW) and the sample size in West Germany (NW) and in East Germany (NE). The running variable is negative for West Germany and positive for East Germany. The specification ‘Baseline’ includes a 5 km border dummy as control, ‘Land cover’ adds the percentage of land cover categories (crop, grassland, forest, water, built up) within the 1 km² sample polygon as control, ‘Field size’ adds additionally the field size as control, ‘Crop diversity’ adds additionally crop diversity as control, ‘EVI’ adds the EVI index as control while ‘Land cover diversity’ includes additionally general land cover patch size and land cover diversity as controls. Standard errors are clustered at the district level.
To further explore the relationship between bird diversity and land cover simplification, we use a cross-sectional regression with the complete geographic range of the CBBS survey.

Our estimation is based on spatial differences of land cover controlling for a range of geographic variables. The baseline regression equation is

\[ Y_{jt} = \alpha_1 LC_{jk} + \alpha_2 FS_{jk} + \alpha_3 CD_{jk} + \alpha_4 EVI_{jk} + \alpha_5 LD_{jk} + X_{jk} + \theta_t + \epsilon_{jkt} \]  

where \( Y \) is our outcome (bird diversity) in sample area \( j \) in district \( k \) and year \( t \), \( LC \) is a vector of land cover shares within the sample areas (cropland, grassland, buildings, water, forest, other) with

**Figure 9** Bird diversity and land cover

Notes: The figure shows the point estimates and the 95% confidence intervals for our cross-sectional regression on bird diversity and land cover characteristics. The “Baseline” specification includes the illustrated variables plus elevation, mean temperatures, mean precipitation levels, and year fixed effects. The “Fixed effects” specification includes additionally district level fixed effects. The “Simpson” specification is as the fixed effects specification but measures bird diversity with the SDI. The “Agriculture” specification is as the “Fixed effects” specification but measures bird diversity as the number of cropland bird species. We standardize all variables such that the magnitude becomes directly comparable. The standard errors of all specifications are clustered at the district level.

5 | LAND USE AND BIRD DIVERSITY

To further explore the relationship between bird diversity and land cover simplification, we use a cross-sectional regression with the complete geographic range of the CBBS survey.

5.1 | Estimation

Our estimation is based on spatial differences of land cover controlling for a range of geographic variables. The baseline regression equation is

\[ Y_{jt} = \alpha_1 LC_{jk} + \alpha_2 FS_{jk} + \alpha_3 CD_{jk} + \alpha_4 EVI_{jk} + \alpha_5 LD_{jk} + X_{jk} + \theta_t + \epsilon_{jkt} \]  

where \( Y \) is our outcome (bird diversity) in sample area \( j \) in district \( k \) and year \( t \), \( LC \) is a vector of land cover shares within the sample areas (cropland, grassland, buildings, water, forest, other) with
forest being the baseline category.\(^9\) FS measures field size, CD is crop diversity, EVI is the EVI index within the sample area, LD is land cover diversity, X is a vector of controls that includes mean precipitation, mean temperature and elevation, \(\theta_i\) are year fixed effects, and \(\epsilon_{jkt}\) is the error term. The errors are potentially correlated across space and within sample sites over time. We therefore use spatial standard errors (Conley, 1999) with a Bartlett Kernel, 25 km distance cutoff and autocorrelation correction across all years. The independent variables may be correlated with other landscape characteristics, which could also affect bird diversity directly and may therefore cause omitted variable bias. We therefore include district-level fixed effects in a robustness test and thus use only differences of land cover characteristics and bird diversity across sample areas within the same districts to estimate our coefficients. All variables are expressed in terms of standard deviations so that the magnitude of the estimates is directly comparable.

5.2 | Results

The baseline results show that bird diversity declines with cropland extent and it increases with land cover diversity (Figure 9). All other land cover types have no significant or consistent impact on bird diversity compared to the natural vegetation in the study area (forests). The results of the baseline specification show that an increase in cropland extent by one standard deviation (29%) reduces bird species richness by 0.15 standard deviations (6 species). Including district fixed effects reduces the estimate to zero. In contrast, crop cover increases bird species richness of cropland bird species as expected. A one standard deviation increase of crop cover increases the diversity of cropland species by 0.73 standard deviations.

The effect of land cover diversity on general species richness is even larger than the impact of crop cover. A one standard deviation increase of land cover diversity increases species richness by 0.42 standard deviations (4.6 species). Including district fixed effects or using the SDI to measure biodiversity has only minor impacts on the estimate. Although landscape diversity is less important for cropland birds, our results indicate that a one standard deviation increase of land cover diversity increases cropland bird species by 0.22 standard deviations.

6 | DISCUSSION

Agriculture is the main driver of global biodiversity loss (Tilman et al., 2017). Underlying the environmental impacts of agriculture are both agricultural expansion and changes within the agricultural landscape. Yet, disentangling the contributions of agricultural expansion from changes in management practices has been stymied by a lack of disaggregated production and biodiversity data collected across the same area and at the same spatial scale. The impact of changes within agricultural landscapes are particularly difficult to study because they are complex, data are often private, and causality between these changes and biodiversity loss is difficult to establish.

In this paper, we create a new data set combining bird observation from opportunistic citizen science and systematic bird surveys with land cover characteristics to study the impact of farm size on bird diversity. We use these data in a regression discontinuity approach that exploits the discontinuous change in farm size at the former inner German border to identify the causal impact of farm size on bird diversity. Although Germany has been unified for more than 30 years, the legacy of former farm collectivization in East Germany is still largely visible. This legacy of large industrial farms in East Germany and relatively small family-owned farm operations in West Germany allows us to compare the environmental outcomes from both regimes at the same time and within the same economic and political environment.

\(^9\)Forest is the natural vegetation in the study area (Bohn & Gollub, 2006).
We find that there is a large and discontinuous decline of local bird diversity at the former inner German border. Matching our bird diversity observations with land cover characteristics allows us to study the mechanisms through which farm size affects bird diversity. Controlling for field size, crop diversity, and land use intensity does not explain the decline in bird diversity at the former inner German border. However, adding land-cover diversity as a control renders the estimated impact of farm size on bird diversity statistically indistinguishable from zero. We interpret this finding as evidence for the importance of the indirect land cover changes associated with increased farm size. Although cropland birds are even more affected by farm size than other species, their response to farm size is not explained by land cover diversity.

Using the CBBS data in a cross-sectional analysis across Germany helps to interpret these results further. The results confirm our previous findings. They show that bird diversity declines with crop-land extent, and it increases with land cover diversity. These results therefore suggest that non-crop habitat plays a crucial role for conserving bird diversity. Our findings therefore highlight the importance of analyzing the agricultural changes in a landscape context.

The results based on the citizen science data (eBird) and based on the structure bird survey (CBBS) are qualitatively similar. However, the results based on the eBird data suffer from several potential biases including site selection bias, species selection bias, endogenous observation effort, and spillovers that can lead to a possible overestimation or underestimation of the impact of farm size on bird diversity. Although we suggest that some of these biases can be mitigated by, for example, including additional controls, they also stress the importance of structured biodiversity monitoring.

In the following we discuss three important caveats. First, farm size has no direct impact on biodiversity. A farm is an economic construct and leaving field size, inputs, management, and so forth unchanged, there are no possible mechanisms through which farm size itself could affect biodiversity or any other outcome. However, larger farms are associated with larger fields (Lesiv et al., 2019; Samberg et al., 2016), less crop and land cover diversity (Ricciardi et al., 2021), a larger share of hired labor (Eastwood et al., 2010), and higher mechanization (Foster & Rosenzweig, 2017; Muyanga & Jayne, 2019; Sheng et al., 2019). These characteristics of larger farms are likely to affect biodiversity. Although the relationship between farm size and these mechanism is well-established (e.g. Ricciardi et al. 2021), their relationship to farm size could have been directly affected by former land use policies. These additional direct impacts of former land use policies on the mechanisms that link farm size to biodiversity in our study limit the external validity of the quantitative estimates while leaving the qualitative results unaffected.

Second, the mechanisms that link farm size to biodiversity include field size, land cover diversity, and land use intensity. We use the EVI as proxy for land use intensity. Measuring land use intensity directly (e.g. the use of agrochemicals) would improve our analysis (see e.g. Li et al. 2020), but to our knowledge, these data are unavailable for Germany. The impact of farm size on biodiversity through pesticides remain therefore largely part of the direct effects (i.e., the residual impact of farm size on bird diversity after controlling for other mechanisms.)

Third, our analysis focuses on common bird species. Rare species such as the great bustard or agricultural specialists such as skylarks have a minor influence on our results. These species, however, may be sensitive to changes in agricultural management (Gayer et al., 2019), but because they are either rare or constitute only a small fraction of local bird diversity, their influence on our results is negligible. Such species are, however, important for bird diversity at a larger scale and may require special protection. Our results mainly concern those species in agricultural landscapes that are common but crucial for the provision of ecosystem services (Kleijn et al., 2015). Our results show that providing a mix of non-crop habitat or different crop types within the agricultural landscape is crucial for biodiversity conservation and can mitigate the negative impact of agricultural industrialization.

11Modeled data such as the PEST-CHEMGRIDS are likely to miss the discontinuity at the border and may introduce omitted variable problems as they use other land use and economic variables to predict pesticide use.
ACKNOWLEDGMENTS
For their helpful comments and discussion, we thank Amy Ando, Patrick Baylis, Julia Brewer, Olivier Deschenes, Eyal Frank, Sumee Gulati, Claire Kremen, Tobias Kuehmerle, Anouch Missirian, Navin Ramankutty, and three anonymous reviewers. We further thank the thousands of volunteers of the German Common Bird Monitoring (CBBS) program as well as all eBird participants, without whom this project would not be possible. CL gratefully acknowledges support by the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No 796451 (FFSize). AL acknowledges the US National Science Foundation (DEB 2042526). The German Common Bird Monitoring (CBBS) is financially supported by the Federal Agency for Nature Conservation (BfN) and the Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (BMU).

REFERENCES


SUPPORTING INFORMATION
Additional supporting information may be found in the online version of the article at the publisher’s website.