DEPARTMENT OF ECONOMICS
JOHANNES KEPLER UNIVERSITY OF LINZ

The causal effect of agritourism on farm survival

by

Matthias FIRGO
Dieter PENNERSTORFER

Working Paper No. 2315
October 2023

Johannes Kepler University of Linz
Department of Economics
Altenberger Strasse 69
A-4040 Linz - Auhof, Austria
www.econ.jku.at

dieter.pennerstorfer@jku.at

The causal effect of agritourism on farm survival*

Matthias Firgo† and Dieter Pennerstorfer‡

September 12, 2023

Abstract

We use two waves of census data with detailed farm characteristics for the total population of Austrian farms to examine the causal effect of agritourism on farm survival. To account for self-selection into agritourism, we use regional variation in tourism intensity that is exogenous to individual farms. On average, agritourism causally increases survival probabilities by 10.3 percentage points over an eleven-year period, which is both large and statistically significant. Marginal effects vary by farm characteristics and are as large as 15.9 percentage points for some sub-populations. Agricultural policies to facilitate entry into agritourism can therefore be effective in keeping farms in the market. Our analysis shows that the magnitude of the estimated coefficients is heavily biased as long as we do not adequately account for endogenous self-selection into agritourism. This suggests that even with a big database, an appropriate identification strategy is required to obtain causal and thus policy-relevant estimates.

Keywords: agritourism, firm survival, agriculture, diversification, big data, bivariate probit

JEL Classification: Z30, Q12, L25

*This work was supported by the Austrian National Bank (OeNB) Anniversary Fund (Project-Number: 17658). The Austrian National Bank did not interfere in any stage of the research.

†HM Munich University of Applied Sciences, Department of Tourism & Austrian Institute of Economic Research (WIFO), Email: `matthias.firgo@hm.edu`.

‡Johannes Kepler University Linz, Department of Economics; Email: `dieter.pennerstorfer@jku.at`.

1 Introduction

Efficiency gains and economies of scale have caused a continuous structural change in the agricultural sector. The number of farms in the EU-27 decreased by 5.3 million holdings (37 %) between 2005 and 2020, while the average farm size increased substantially over the same period (Eurostat, 2023). In these difficult circumstances, farms have several options to remain profitable and survive in the market. The most obvious strategies are to increase the size of the farm to achieve a more efficient scale of production or to "deepen" (Van der Ploeg et al., 2003) farming activities to produce higher value-added goods (by improving product quality or switching to organic farming, for example). Alternatively, farmers can take advantage of non-agricultural income sources either by finding off-farm employment opportunities (thus switching to part-time farming) or by producing non-agricultural goods and services on the farm, such as offering farm-stay accommodation (agritourism).

The relation between off-farm employment opportunities, off-farm employment and part-time farming are theoretically unclear and empirical results are mixed.[1] On the one hand, farmers may pursue non-financial goals (such as farming lifestyle) and could see off-farm employment as a way to stay in the market due to cross-subsidization of the farming business by off-farm income (Goddard, 1993, Kimhi, 2000, Zimmermann and Heckelei, 2012). On the other hand, good off-farm employment opportunities could increase the opportunity costs of labor (Hallam, 1991, Zimmermann and Heckelei, 2012), and part-time farming might be only an intermediate step in leaving the agricultural market entirely.

Agritourism, on the other hand, is an easy-to-provide service that generates additional on-farm income and, by definition, requires an active farm. Both aspects—additional income and the requirement of an active farm—could therefore prevent farms from exiting the market. However, research on the economics of agritourism remains surprisingly thin (Schilling et al., 2014, p. 69). Considering that agritourism is an easily exploitable source of income directly linked to farm profitability, this limited interest is rather surprising. This is even more so given both the enormous

[1] At the regional level, Goetz and Debertin (2001) find that off-farm employment does not influence the net rate of farm exits when analyzing U. S. county-level data, but report more heterogeneous results when analyzing sub-samples. At the farm level, empirical evidence shows that part-time farmers have higher exit probabilities in Germany (Pfeffer, 1989), the U. S. (Roe, 1995) and Austria (Weiss, 1996, 1999), whereas Kimhi and Bollman (1999) find the opposite results for both Canada and Israel.

political and financial efforts to preserve the small-scale agricultural structure in Europe and the economic importance of tourism, especially in many rural and lagging regions.

Against this background, the main contribution of this article is twofold. First, we contribute to the scarce empirical literature examining the influence of agritourism on farm performance. Second, we illustrate that even with a big database, an appropriate identification strategy based on economic theory is required to obtain meaningful and thus policy-relevant results (cf. Mullainathan and Spiess, 2017). Our empirical analysis is based on the full population of about 220,000 farms in Austria. The comprehensive micro-level data include a wide range of farm and farmer characteristics. Nevertheless, we show that relying solely on the large number of observations and potential control variables in a simple regression framework leads to highly significant and robust, but completely inconsistent results. The parameter estimates change substantially once we consider that farms' decisions to offer agritourism are endogenous. Controlling for this self-selection in an instrumental variables framework allows us to assess the causal effect of agritourism on farm survival, which is indeed the (policy relevant) parameter of interest.

With respect to the first contribution, the largest part of this literature evaluates the charcteristics (Jeczmyk et al., 2015), the determinants (Bagi and Reeder, 2012, Joo et al., 2013, Khanal and Mishra, 2014, Lupi et al., 2017, Yeboah et al., 2017), or the motives (McGehee and Kim, 2004, Nickerson et al., 2001) of farms to engage in agritourism.[2] These studies focus mainly on individual (farm) characteristics determining the farms decisions to provide farm-stay accommodation and find, among other things, that farmers' education and age have a significant influence (see, e.g., Joo et al., 2013). Whether farms are located in touristic regions (and can expect high demand for their lodging services) is treated surprisingly superficially.[3] A systematic analysis of the relationship between local tourism demand and farms' decisions to offer farm-stay accommodation, however, is not available. Only a few articles (Joo et al., 2013, Khanal and Mishra, 2014, Schilling et al., 2014) examine the influence of agritourism on farm performance, all of which use data from U. S. farms.[4]

[2]A different perspective is taken by Carpio et al. (2008), Hill et al. (2014), Ohe and Ciani (2011, 2012) and Santeramo (2015), who investigate demand for (or demand-related aspects of) agritourism.

[3]The empirical literature documents that the landscape around the farmsteads (Lupi et al., 2017), whether the farms are located near cities (Yeboah et al., 2017) and close to land enrolled in conservation programs (Bagi and Reeder, 2012), as well regional dummy variables (e.g., Lupi et al., 2017)—variables that could be interpreted as indicators of tourism demand—influence the farms' decisions to engage in agritourism.

[4]Schilling et al. (2014) find significantly positive effects of agritourism on farm performance. Joo et al. (2013), on the other hand, document a positive influence only for small farms, but could not find a statistically significant relationship either for large farms or for the entire sample. Khanal and Mishra (2014) analyze a slightly different

Concerning the second contribution, we offer an alternative (and a more plausible) empirical approach than the existing literature that has relied on a selectivity correction method (Khanal and Mishra, 2014) or on matching techniques (Joo et al., 2013, Schilling et al., 2014). We use an instrumental variables (IV) approach based on a recursive bivariate probit model to address the (endogenous) selection of farms to engage in agritourism and are thus able to evaluate the causal effects of participating in agritourism on farm survival.[5] The instrumental variables proxy local demand for tourism services and are available at a fine spatial scale (municipality level). Local demand differs substantially across regions, has a large impact on farms' decision to offer farm-stay accommodation, and is plausibly exogenous to farmer heterogeneity in unobserved characteristics. This framework allows us to estimate the causal effects of agritourism on farm performance.

Additionally, the article adds to the existing literature by drawing on detailed data on the importance of tourism at a highly disaggregated regional level, which is novel to the existing literature on the determinants of supply in agritourism. Further, we are the first to provide evidence on the effect of agritourism engagement at the intensive margin (the number of beds for rent) in addition to the extensive margin. Finally, our data differs substantially from other empirical contributions estimating the effects of agritourism on farm performance: We use farm survival as an indicator of revealed competitiveness that can be observed easily and that enables us to utilize the full population of farms over a long time period (11 years). In contrast to the literature, which focuses on the U. S., we provide the first empirical evidence for a country in Europe, where the agricultural sector is characterized to a much greater extent by small-scale farming.

Our results show that agritourism causally increases survival probabilities by more than 10 percentage points on average over our observation period (of 11 years), with up to 15.9 points for some sub-populations. The IV estimates differ substantially from results derived from simple probit regressions (which underestimate the effect by about 50 %), suggesting that the self-selection bias is

---

research question and evaluate the impact of two diversification decisions simultaneously (offering agritourism and taking up off-farm work) using a sample of small farms. They conclude that household income increases most when farms choose to diversify in both dimensions rather than just one.

[5]The crucial assumptions necessary to perform matching methods are that selection into treatment (i.e., engaging in agritourism) depends on observable variables only, while there are no unobservable variables that affect both farms' decisions to offer farm-stay accommodation and their survival probabilities (see Cameron and Trivedi, 2005, for details). However, it seems plausible that unobservable variables such as a farmer's ability, effort, managerial or entrepreneurial skills influence both selection (into agritorism) as well as outcome (farm survival). Unlike matching techniques, the IV-approach pursued in this article can eliminate the bias due to self-selection into treatment even in this case. See Miller (2021), among others, for a critical article on the limited scope of matching techniques to address the non-random selection into treatment.

large and endogeneity concerns have to be addressed carefully. Results on the intensive margin point towards the existence of an inverse U-shaped effect, suggesting that agricultural policies supporting entry into agritourism should therefore target moderate numbers of beds. Given the big micro-level data set (in terms of the number of observations and control variables), our results are robust to a number of different specifications and estimators.

The remainder of this article is organized as follows: Section 2 discusses the research design and the identification strategy. In Section 3, we present the data used in the empirical analysis. Section 4 reports and discusses the main results and the sensitivity analysis. Section 5 concludes.

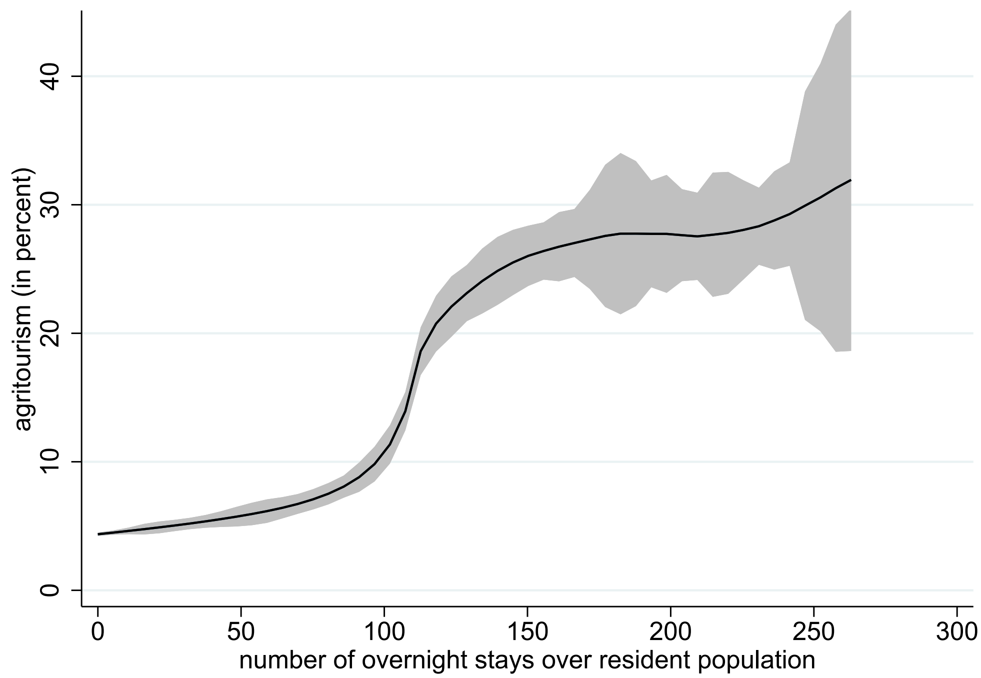
2 Empirical approach and identification strategy

This article addresses the impact of agritourism on farm survival and evaluates whether agritourism contributes to slow down the structural change in this industry. Throughout the paper we define agritourism as active farms providing farm-stay accommodation. We first provide stylized facts about the relationship between the intensity of local tourism, farm decisions to offer lodging services, and farm survival that guide our analysis. This is followed by a formal discussion of our empirical model.

The decision of farms to offer farm-stay accommodation is closely linked to local demand for this kind of tourist services. Descriptive evidence, provided in Figure 1 for Austrian municipalities, supports this notion: In regions with low tourism intensity (less than 50 overnight stays relative to the resident population per year) only 5 % of all farms offer farm-stay accommodation. The share of farms offering agritourism increases sharply at an intensity of around 100 overnight stays per capita, and more than a quarter of all farms rent out rooms and apartments in municipalities with a tourism intensity of more than 150 overnight stays. It is important to note that tourism intensity varies at a small spatial level and that only a quarter of the variation in intensity at the municipality level can be explained by variation *between* districts (the next higher regional unit), while most of the variation is between municipalities *within* districts.

Figure 2 shows the relationship between the level of tourism and farm survival rates at the municipality level. Panel a) shows a scatterplot and a linear regression between these two variables, indicating that survival rates are higher in regions with intense tourism. The positive correlation is

Figure 1: Correlation between local tourism intensity and agritourism

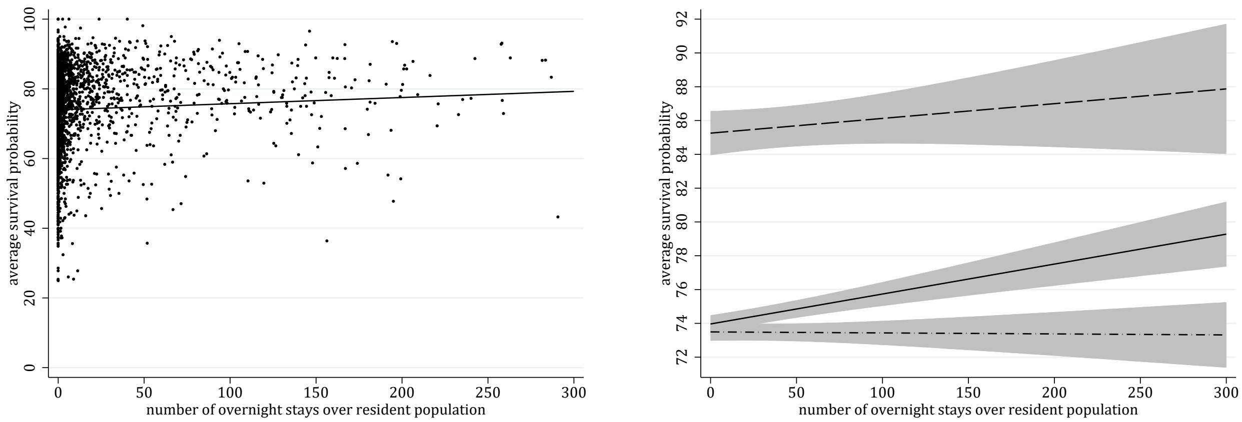


Notes: The figure shows the non-parametric relationship between the share of farms offering farm-stay accommodation and the ratio of overnight stays to resident population in 1999 at the municipality level. The local polynomial smooth is based on an Epanechnikov kernel with a polynomial smooth degree of 0 and a bandwidth of 50.

even more evident in Panel b), where we narrow the range of the survival probabilities (the vertical axis). This figure suggests that an increase in overnight stays per capita by 100 is associated with 1.8 percentage points higher farm survival rates. When we divide farmers into two groups depending on whether they offer farm-stay accommodation or not, we find substantial differences between these groups: The survival rate of farms offering farm-stay accommodation is 85.6 %, which is about 12.1 percentage points higher than for farms without agritourism. However, the survival rates of each of the two groups of farms are not significantly related to the intensity of tourism in the region, as illustrated in Panel b) of Figure 2. The higher survival probabilities in areas with intense tourism seem to come only from a larger share of farmers offering farm-stay accommodation, suggesting that the product diversification associated with agritourism may indeed increase farms' competitiveness and slow down structural change by allowing a larger share of farms to stay in the market.

Figure 2: Correlation between local tourism intensity and farm survival

a) All farms   b) Farms with and without agritourism



Notes: The figures indicate average farm survival rates (in percent) between 1999 and 2010 and the ratio of overnight stays to resident population in 1999 at the municipality level. The straight lines in panels a) and b) show the prediction of average farm survival rates from a linear regression of survival rates on the ratio of overnight stays to resident population at the municipality level. The dashed line and the dashed-dotted line in panel b) indicate predictions of the average farm survival rates for farms offering farm-stay accommodation (dashed line) and for farms not offering farm-stay accommodation (dashed-dotted line). Gray areas illustrate the 95 % confidence intervals.

Farm survival (or, conversely, farm exit) between to points in time is usually estimated using conventional probit or logit models. In our case, the profits of farm i (captured by the latent variable $y^*_{1,i}$) are unobservable and depend on observable farm and farmer characteristics $\boldsymbol{x}_i$ as well as on the farm's decision whether to offer farm-stay accommodation, which is indicated by the dummy variable $y_{2,i}$. The model can be summarized by the following Equation (1):

$$y_{1,i} = 1 \text{ if } y^*_{1,i} = \alpha_1 y_{2,i} + \boldsymbol{x}_i\boldsymbol{\beta}_1 + \epsilon_{1,i} > 0, \tag{1}$$

with α_1 and $\boldsymbol{\beta_1}$ as the (vector of) parameters and $\epsilon_{1,i}$ as the error term. We only observe whether a farm stays in the market (which means that $y_{1,i} = 1$), which is the case if the latent variable $y^*_{1,i} > 0$, i.e., if profits (including subsidies) of farm i are positive. Therefore, rather than estimating the (unobserved) latent variable $y^*_{1,i}$ in Equation (1) directly, we estimate the probability whether a farm survives, i.e., $\text{Prob}(y_{1,i} = 1|y_{2,i}, \boldsymbol{x}_i)$.

However, when estimating farm survival probabilities, it is problematic to simply include farms' decision to participate in agritourism as an additional explanatory variable. As discussed above and highlighted in Figure 1, this may lead to self-selection bias because this decision is not exogenous. Whether a farm offers agritourism ($y_2 = 1$) depends on the unobserved latent variable $y^*_{2,i}$, indicating the additional farm profits when offering lodging services. These additional profits again depend on farm and farmer characteristics $\boldsymbol{x}_i$, as well as on the local demand for overnight stays in municipality m, captured by the vector $\boldsymbol{z}_m$, as summarized in Equation (2):

$$y_{2,i} = 1 \text{ if } y^*_{2,i} = \boldsymbol{x}_i\boldsymbol{\beta}_2 + \boldsymbol{z}_m\boldsymbol{\gamma}_2 + \epsilon_{2,i} > 0. \tag{2}$$

$\boldsymbol{\beta}_2$ and $\boldsymbol{\gamma}_2$ are the vectors of parameters and $\epsilon_{2,i}$ is the error term. Again, we do not observe the latent variable $y^*_{2,i}$, but only whether a farm rents out rooms and apartments (which means that $y_{2,i} = 1$), which is the case if the latent variable $y^*_{2,i} > 0$, i.e., if offering farm-stay accommodation generates additional profits.

Due to potential self-selection bias, the correlation of the error terms, $corr(\epsilon_{1,i}, \epsilon_{2,i})$, may be different from zero. A simple probit estimation on farm survival will therefore give inconsistent parameter estimates for α_1 and $\boldsymbol{\beta}_1$. We thus estimate a recursive ("triangular") bivariate probit model (see Greene 2011, p. 786, and Coban 2020, p. 23, for details). With this approach, the binary variable on whether farm i offers agritourism, $y_{2,i}$, is treated as an endogenous variable when estimating the probability that a farm survives, $\text{Prob}(y_{1,i} = 1|y_{2,i}, \boldsymbol{x}_i)$. As long as the error terms $(\epsilon_{1,i}, \epsilon_{2,i})$ follow a bivariate normal distribution, a bivariate probit model can be estimated using maximum-likelihood techniques.[6] The recursive bivariate probit model is properly identified if there is at least one variable available that is correlated with a farm's decision to provide farm-stay accommodation, but uncorrelated with a farm's probability to exit the market. Such instrumental variables are captured by matrix $\boldsymbol{z}_m$ in Equation (2). We instrument the dummy variable $y_{2,i}$ with several measures on tourism intensity at the level of municipalities m. Since the instrumental variables are calculated at a higher regional level (municipality) than the observation units (the farms), we cluster the residuals at the municipality level.

We are confident that variables capturing the intensity of local tourism are strong and valid instruments. First and foremost, we observe a high correlation between local tourism intensity and the probability that farms offer farm-stay accommodation (see Figure 1). On the other hand, descriptive evidence suggests that the intensity of local tourism does not seem to influence farm

[6]See Wooldridge (2001, Chapter 15) for an introduction, Lewbel et al. (2012) for a recent and comprehensive treatment on probit models with endogenous binary regressors (in particular on the MLE-approach, pp. 816 f.), and Evans and Schwab (1995) and Coban (2020) for empirical applications.

survival through channels other than agritourism supply (see Figure 2). One may argue, for example, that the tourism industry could influence farmers' decisions to exit the market by providing employment opportunities. This could make working in agriculture less attractive or, contrariwise, could enable part-time farming due to the availability of part-time jobs in tourism. However, labor market areas are typically much larger than municipalities. Dummy variables at the district level, regions that correspond much better to labor market areas, allow us to control for potential labor market opportunities of tourism. We can thus adequately control for this alternative channel. Furthermore, we provide statistical tests on the strength and the validity of our instruments: As a sensitivity analysis, we estimate Equation (1) and Equation (2) as a linear probability model (LPM) via 2SLS, which allows us to provide test statistics on the exclusion restriction and on the quality of the instruments.

3 Data

To evaluate the impact of agritourism on farm survival, we use data at different levels of aggregation from two different sources: First, we use data from the Farm Structure Surveys (FSS), which provide information on the entire population of Austrian farms ("agricultural holdings"). The Farm Structure Surveys provide micro data at the farm level of excellent quality. These data are collected by the Austrian Statistical Office ("Statistics Austria") and provided by the Federal Ministry of Agriculture, Forestry, Environment and Water Management (BMLFUW) for this project.[7] A general census is conducted approximately every ten years, and we use data from 1999 and 2010.[8] Second, we use data from Statistics Austria's "Accommodation Statistics" for the year 1999, which provides information on tourism at the regional (municipality) level. This information is used to construct instrumental variables for the regression analysis (as captured by matrix $\boldsymbol{z}_m$ in Equation (2)).

[7]The BMLFUW promotes empirical research by offering access to these data as a benefit in kind. The data can be used for scientific purposes, but researchers are not allowed to disclose individual data and are obliged to make their results publicly available.

[8]In addition to census data, more frequent data are available for representative (stratified) samples of about 30,000 farms collected in 2003, 2005, 2007, 2013 and 2016. However, for our analysis of survival rates, we need data on the entire population of farms, since the sample data cannot be used to extract information on market exits, as described below.

3.1 Farm-level data

Data from the FSS include detailed characteristics at the farm level (economic and spatial size, soil quality, labor force, output by product, type of farming, farm manager characteristics, topography such as difficulty of terrain or alpine pastures, eligibility for specific farm-level subsidies), local geographic information (municipality characteristics, municipality code), and a unique farm-level identifier that remains constant over time. Summary statistics and a detailed discussion of all variables included in the analysis can be found in Appendix A.1. The FSS data for Austria also contain detailed information on agritourism, namely whether farms offer lodging services and (if so) how many beds they rent out.

Survival and *Agritourism* are the two dependent variables of the bivariate probit model, corresponding to the variables $y_{1,i}$ and $y_{2,i}$ in Equation (1) and Equation (2), respectively. $Survival = 1$ if the unique farm identifier is found in the 1999 and 2010 FSS data, which means that a farm that was present in 1999 is still active in 2010. Farms exiting between 1999 and 2010 ($Survival = 0$) are identified if they appear in the 1999 FSS data but not in the 2010 data. $Agritourism = 1$ if farms rent out tourist accommodation according to the 1999 FSS data, and $Agritourism = 0$ otherwise. As shown in Table 1, 74.7 % of all farms active in 1999 survived in the market until 2010, which means that about a quarter of Austrian farms left the market. Nearly 15,500 or 7.2 % of all farms participated in agritourism in 1999. For these, the survival probability was 88.0 %, and thus substantially and significantly[9] higher than for the farms that did not offer farm-stay accommodation (with a survival probability of 73.6 %).

Table 1: Farm survival and agritourism

	Agritourism		
Survival	Yes	No	Total
Yes	13,608	147,766	161,374
No	1,850	52,896	54,744
Total	15,458	200,660	216,118
Survival probability	88.0 %	73.6 %	74.7 %

Note: The figures are based on data from the Farm Structure Surveys (FSS) from 1999 and 2010.

[9]The difference is significant at the 5 % level according to a two-sample t-test with unequal variances.

3.2 Regional data on tourism

To derive indicators of local demand for lodging services, we use regional tourism data from 1999. These data contain information on the capacity and volume of tourism services at the municipality level. The information on capacity includes the number and the type of lodging establishments (hotels, motels, camping sites, etc.) and the number of beds (as a measure of the capacity) offered by each type.[10] Additionally, these data include the number of overnight stays for each type of accommodation. For privacy reasons, data are not reported if the number of overnight stays in the municipality is less than 1,000 or if there are fewer than three lodging establishments.

Based on these data, we calculate three variables to measure the intensity of tourism. First, we use a dummy variable ($LowTourism$) that equals one if the number of overnight stays is not reported (and zero otherwise) because there are fewer than 1,000 overnight stays per year and/or fewer than three lodging establishments. As summarized in Table 2, this is the case for about 34 % of all municipalities in 1999. In turn, 29.6 % of all farms were located in municipalities with low tourism.[11] Second, we calculate the number of overnight stays per resident in the municipality ($OvernightStays$). This variable includes stays by domestic and international tourists for all kinds of touristic purposes (recreation, business, etc.). The average number of overnight stays per capita is 24.70 (including zeros), but the distribution of this variable is highly right-skewed with considerable dispersion and a maximum of 1,071.79. The third tourism intensity variable is the average number of beds per lodging establishment ($BedsPerFacility$). It is calculated as the number of tourist beds in relation to the number of accommodation establishments in a municipality. The average number of beds per accommodation facility provides important information on the local structure of tourism supply.[12] The variable is again right-skewed with a mean of 13.72 and a maximum of 1,388. Both $OvernightStays$ and $BedsPerFacility$ are unobservable for municipalities with low tourism (i.e., when $LowTourism = 1$). These missing observations are replaced by zeros.[13]

[10] Unfortunately, farms offering lodging services are not a separate category. This is not a major problem in the empirical analysis, as agritourism represents only a very small fraction of the tourism industry.

[11] Although there is no precise information on tourism intensity for a sizable share of farms, we do know (because of the low threshold of 1,000 overnight stays per year) that tourism intensity in the regions of these farms is virtually zero.

[12] With regard to the structure of local accommodation supply, two countervailing effects can be expected: Smaller establishments are probably better substitutes for agritourism accommodation (competition effect). On the other hand, if tourists in a given region have particular preferences, or if small lodging establishments (either farms or other facilities) fit better into the landscape than large hotels or resorts, a higher proportion of smaller establishments could

Table 2: Descriptive statistics of variables on tourism at the municipality level

Variable	Mean	Std. Dev.	Min	Max	N
LowTourism	0.34	0.47	0	1.00	2,358
OvernightStays	24.70	72.34	0	1,071.79	2,358
BedsPerFacility	13.72	34.00	0	1,388.00	2,358

Note: The data were collected by Statistics Austria in 1999 as part of the "Accommodation Statistics".

In addition to farm and farmer characteristics, (off-farm) labor market opportunities and sociodemographic characteristics of the resident population may also affect the probability of farm survival. To account for these and other unobserved region-specific characteristics, district-level fixed effects are included in the main specifications. In Austria, the district level corresponds to the level between the higher NUTS 3 and the lower municipality (LAU) level.[14] Due to the small number of farms and the low prevalence of agritourism, 15 purely urban districts were merged with their respective surrounding districts. In addition, the 898 farms within the city of Vienna are excluded from the analysis due to the low prevalence of agritourism (only 5 out of 898 farms offer lodging services) and the substantial heterogeneity in socioeconomic and structural characteristics to neighboring districts.[15]

4 Results

In discussing the results, we first provide descriptive evidence from simple probit models in Section 4.1. A probit model regressing farm *Survival* on participation in *Agritourism* provides biased results if a farm's self-selection into *Agritourism* depends on unobservable characteristics that also

be positively correlated with agritourism (composition effect).

[13] In the empirical analysis, we include the logarithmic transformation of *OvernightStays* and *BedsPerFacility*. Instead of adding a small number before the log transformation, we replace missing observations after the logarithmic transformation with zeros, as *LowTourism* serves as a missing flag and thus identifies the zeros for missing data. Alternatively, we also estimated all models adding one to the levels of overnight stays (including to the zeros for missings) before the log transformation instead of adding zeros for missing data afterwards. The results are virtually unchanged (and are available from the authors upon request).

[14] The Nomenclature of Territorial Units for Statistics (NUTS) is a geocode standard for referencing the subdivisions of countries for statistical purposes. It is developed and regulated by the European Union and covers the member states of the EU. In 1999, Austria (NUTS 0) consisted of 3 NUTS 1 (groups of federal states), 9 NUTS 2 (federal states), 35 NUTS 3 (groups of districts) and 2,359 LAU (municipalities) regions. The number of districts was 99 if the capital Vienna is considered as a single district.

[15] We therefore include 83 instead of 99 district fixed effects. Vienna and Viennese farms are already excluded from the descriptive statistics in Table 1 and Table 2.

affect the probability of farm *Survival*. In this case, the error terms of Equation (1) and Equation (2) are correlated, and the parameter estimates for *Agritourism* cannot be interpreted causally. Therefore, in the subsequent Section 4.2, we present and discuss the estimation results based on recursive bivariate probit models that take into account the (possibly endogenous) self-selection of farms into agritourism. To show that the results are not driven by the specific estimation procedure, we provide results on linear probability models (LPMs) in Section 4.3. An advantage of LPMs is that they allow us to perform statistical tests on the validity and strength of our instruments. Finally, in Section 4.4, we provide results on the intensive margin to evaluate whether the bed capacity of agritourism farms also affects their performance.

4.1 Descriptive evidence: Simple probit models

Table 3 shows the marginal effects of *Agritourism* on farm *Survival* in a simple probit framework. The baseline specification (1) does not include farm characteristics or regional fixed effects. Model (2) considers a large number of farm characteristics to control for observable heterogeneity that affects a farm's competitiveness and thus its survival probability. Specifications (3) and (4) extend model (2) by adding fixed effects at the district and municipality levels, respectively. For brevity, the parameter estimates for all farm characteristics are relegated to Table B1 in Appendix B. The simple regression in column (1) suggests that *Survival* probabilities are about 17.3 percentage points (pp) higher for firms engaging in *Agritourism*. The effect is reduced to 4.8 pp when detailed farm-level characteristics are included according to specification (2). The size of the marginal effects remains remarkably stable when fixed effects at the district (4.9 pp) or municipality (4.8 pp) level are included, as indicated in columns (3) and (4).[16] The marginal effects of *Agritourism* are signifcantly positive at the 0.1 % significance level for all model specifications reported in Table 3.

4.2 Causal evidence: Recursive bivariate probit models

While the simple probit models suggest a stable but rather moderate effect of agritourism on farm survival of around 5 pp (once we control for observable farm characteristics), these results may be biased because the endogeneity of farms' decisions to participate in agritourism is not

[16]Note the the number of observations drops by 65 in specification (4) because farm *Survival* is perfectly predicted in some municipalities due to municipality fixed effects.

Table 3: Marginal effects of simple probit models

Dependent variable	Model (1) *Survival*	Model (2) *Survival*	Model (3) *Survival*	Model (4) *Survival*
Agritourism	0.173***	0.048***	0.048***	0.047***
	(0.006)	(0.005)	(0.005)	(0.005)
Farm heterogeneity	No	Yes	Yes	Yes
Regional fixed effects	No	No	District	Municipality
Observations	216,118	216,118	216,118	216,053

Notes: Marginal effects are calculated at means. Standard errors are reported in parentheses and are clustered at the municipality level. *** significant at 0.1 %, ** significant at 1 %, * significant at 5 % level. Farm heterogeneity includes 24 variables on farm characteristics and 32 detailed farm type fixed effects. Parameter estimates for all farm characteristics are reported in Table B1 in Appendix B.

accounted for. Therefore, we extend the probit model by switching to the recursive bivariate probit model described in Section 2. The marginal effects of the most important variables explaining participation in *Agritourism* and farm *Survival* are summarized in Table 4. Parameter estimates for all control variables for both regressions are relegated to Table A2 in Appendix A.2 for brevity and are accompanied by a detailed discussion.

Again, specification (1) includes *Agritourism* as the only regressor in the *Survival* equation, while municipality-level tourism data (as described in Section 3.2) are the only explanatory variables in the *Aritourism* equation. The resulting marginal effect of *Agritourism* on *Survival* is 19.6 percentage points. This coefficient is more than halved to 8.7 pp when farm-level characteristics are added to both equations of the recursive bivariate probit model, as done in model (2). The size of the marginal effect is again robust to the inclusion of regional fixed effects and changes only by a small and statistically insignificant amount to 10.3 pp when district fixed effects are added as in specification (3).[17]

The results of the identifying variables on local tourism in the *Agritourism* equation are shown to have strong explanatory power and have the expected signs. Being located in a *LowTourism* municipality reduces the probability for farms to engage in *Agritourism* by about 2.5 pp in the

[17]Given that the tourism intensity variables used as instruments to identify the recursive bivariate probit model vary at the municipality rather than the farm level, we cannot estimate a specification using municipality fixed effects corresponding to model (4) in Table 3. Due to the strong robustness of the marginal effect of *Agritourism* in both Tables 4 and 3 with respect to the inclusion of (different levels of) regional fixed effects, we are confident that the higher level district fixed effects are sufficient to capture any potentially remaining spatial patterns that are not controlled for by the large set of control variables.

Table 4: Marginal effects of recursive bivariate probit models

	Model (1)		Model (2)		Model (3)	
Dependent variable	*Survival*	*Agri-tourism*	*Survival*	*Agri-tourism*	*Survival*	*Agri-tourism*
Agritourism	0.196***		0.087***		0.103***	
	(0.012)		(0.012)		(0.011)	
LowTourism		−0.062***		−0.034***		−0.025***
		(0.003)		(0.003)		(0.003)
ln(OvernightStays)		0.033***		0.026***		0.021***
		(0.001)		(0.001)		(0.001)
ln(BedsPerFacility)		−0.033***		−0.022***		−0.016***
		(0.002)		(0.001)		(0.001)
Farm heterogeneity	No	No	Yes	Yes	Yes	Yes
Regional fixed effects	No	No	No	No	District	District
Observations	216,118		216,118		216,118	
ρ	−0.186***		−0.131***		−0.176***	

Notes: Marginal effects are calculated at means based on the Stata command *rbiprobit* (Coban, 2020). Standard errors are reported in parentheses and are clustered at the municipality level. The parameter ρ denotes the correlation of the error terms. Significance levels for ρ are based on a Wald test with the null hypothesis $\rho = 0$. *** significant at 0.1 %, ** significant at 1 %, * significant at 5 % level. Farm heterogeneity includes 24 variables on farm characteristics and 32 detailed farm type fixed effects. Parameter estimates for all farm characteristics are reported in Table A2 in Appendix A.2.

preferred specification (3). An increase in the tourism intensity measured by the *OvernightStays* per resident by one percent increases the probability for *Agritourism* by 0.021 pp. An increase in the average number of *BedsPerFacility* in tourism establishments by one percent, in turn, reduces the probability for *Agritourism* = 1 by about 0.016 pp. Thus, farms are more likely to provide tourist accommodation in municipalities with smaller rather than large-scale accommodation facilities.

The importance of estimating a recursive bivariate probit model rather than a simple probit model to determine the impact of *Agritourism* on the probability of farm *Survival* is emphasized by two results. First, the marginal effects in the simple probit models are less than 5 percentage points, about half the size of the corresponding effects based on the recursive probit models (when farm characteristics are taken into account). This comparison shows that the bias due to self-selection of farms into *Agritourism* in the simple probit model is substantial. Second, the correlation of the error terms of the two equations in the bivariate probit model, given as parameter ρ in Table 4, is strongly (between −0.19 and −0.13, depending on the specification) and statistically significantly negative.[18]

[18]The null hypothesis of no correlation is rejected by a Wald test at the 0.1% significance level.

The negative values for the ρ parameters indicate that the unobservable variables that increase the survival probability decrease the probability of participation in agritourism. This implies that not accounting for self-selection of farms with (ceteris paribus) lower survival probabilities into agritourism should lead to an underestimation of the effect of *Agritourism* on farm *Survival*, which is exactly what we find when we compare the results of simple with bivarate probit models.

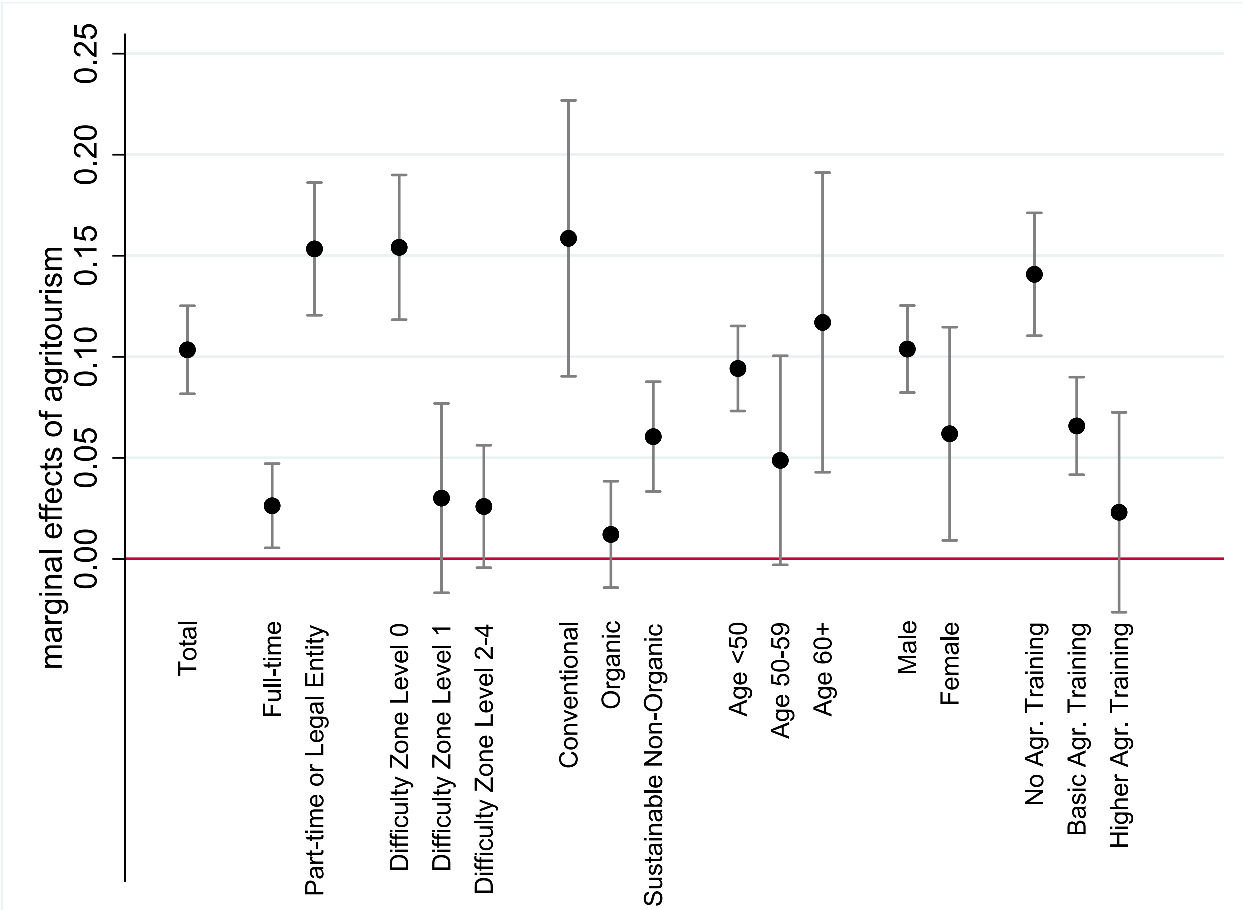
Although we control for a large number of variables (namely 24) to account for farm and farmer characteristics, and also include 32 dummy variables to capture different farm types and 83 district fixed effects, the marginal effects of *Agritourism* are restricted to be homogeneous across all farms. To relax this strong assumption and examine the potential heterogeneity of the effect of *Agritourism*, we re-ran specification (3) of the recursive bivariate probit model reported in Table 4 for different sub-populations of farms with respect to key farm and farm manager characteristics. The results of this analysis are shown in Figure 3. In this diagram, the dots denote the point estimates of the marginal effects of participation in *Agritourism* on farm *Survival*, while and vertical bars illustrate the 95 % confidence intervals.

The effect estimated for the total population is given as a benchmark on the far left of the figure. There is a substantial difference in the effect between full-time and part-time farmers.[19] The positive effect is much larger for part-time farms (+15.3 pp) than for full-time farms (+2.6 pp). Another striking result is the difference in effects for farms in normal vs. difficult (steep, mountainous) terrain: Only farms in locations that are not considered difficulty zones are more likely to survive if they engage in agritourism. This result may seem surprising at first glance. However, the more difficult the terrain, the higher the so-called federal "hardship allowance" and the better the access to structural subsidies for farms in disadvantaged locations and terrain. Therefore, farms located outside difficulty zones are more likely to depend on market income.

Similarly, the results in Figure 3 show that conventional non-organic farms experience the largest positive effect of *Agritourism* on farm *Survival* (+15.9 pp). In contrast, the effect is insignificant for organic farms and relatively small for other sustainable (non-organic) farm types (+6.0 pp). While there is no evidence for differences in effects between age groups and gender of the farm manager, *Agritourism* appears to be a particularly effective strategy for farms managed by farmers

[19]Farms operating as legal entities are grouped with part-time farms, as only a small share (3.6 %) of farms operate as legal entities.

Figure 3: Marginal effects for sub-populations



Notes: The figure illustrates the marginal effects of agritourism on farm survival along with the 95 % confidence intervals. The estimates are based on a recursive bivariate probit model that controls for farm heterogeneity and district fixed effects. The first marginal effect ("Total") corresponds to specification (3) in Table 4. All other marginal effects are based on regressions with the same explanatory variables for the corresponding sub-populations.

without formal agricultural training. The corresponding effect (+14.1 pp) is more than twice as high as for farms with basic agricultural training of the farm manager (+6.5 pp). For farms with higher agricultural training of the farm manager, the effect is small (+2.3 pp) and not significantly different from zero. These results correspond well with the interpretation of the negative sign of the estimated correlation of the error terms in the bivariate probit model, captured by the parameter ρ: Offering lodging services can be a relatively easy source of income for farms that are struggling to operate profitably, which is especially true for conventional part-time farmers and farmers without formal agricultural training (i.e., farms with a high exit probability).

4.3 Sensitivity analysis: Linear probability models

As an alternative to the recursive bivariate probit model, we estimate Equation (1) and Equation (2) by a linear probability model (LPM) in this section. This is to show, on the one hand, that our results do not depend on the specific functional form of the relationship of interest as assumed by the chosen method. On the other hand, this framework allows us to provide statistical tests on the quality of the instruments used. This model estimates participation in *Agritourism* as the first stage equation and farm *Survival* as the second stage equation in the form of a linear 2SLS framework. With this approach, the parameter estimates can be interpreted directly as marginal effects. Results for the three main model specifications (analogous to Table 4) are presented in Table 5, while parameter estimates for all explanatory variables are again relegated to Appendix B (see Table B2).

The results show, first, that the marginal effect of *Agritourism* on farm *Survival* is again significantly positive and of a similar magnitude as in the recursive bivariate probit model (about 9 to 10 pp in specifications (2) and (3)). Second, the instruments are strong according to the test statistics for underidentification (Kleinbergen-Paap rk LM statistic rejects the null hypothesis of underidentification at the 0.1 % level) and weak identification (Kleinbergen-Paap rk Wald F statistic of at least 187). Third, the instruments are valid according to the Hansen J statistic for overidentification: The test does not reject the null hypothesis that the instruments are uncorrelated with the error term at the 5 % level in the preferred specification (3). Therefore, the LPMs confirm a marginal effect of *Agritourism* on farm *Survival* of about 10 pp (which is twice as large as that of the simple probit models from Section 4.1) and support the causal interpretation of the results.

4.4 The intensive margin of agritourism

The analysis so far has been devoted to examining the causal effect of participation in *Agritourism* on farm *Survival* at the extensive margin. While the analysis has found overwhelming evidence of a substantial average effect, some heterogeneity of this effect at the intensive margin also seems reasonable. On the one hand, because of the fixed costs of providing accommodations, it seems plausible that there is a non-linear relationship between profitability and the number of beds provided. Thus, as the number of beds increases, increasingly positive effects on the farm survival

Table 5: Parameter estimates of linear probability models

	Model (1)		Model (2)		Model (3)	
Endogenous variable	*Survival*	*Agri-tourism*	*Survival*	*Agri-tourism*	*Survival*	*Agri-tourism*
Agritourism	0.276***		0.103***		0.090***	
	(0.027)		(0.019)		(0.025)	
LowTourism		−0.047***		−0.027**		−0.018*
		(0.009)		(0.009)		(0.008)
OvernightStays		0.072***		0.070***		0.056***
		(0.002)		(0.002)		(0.003)
BedsPerFacility		−0.044***		−0.036***		−0.028***
		(0.003)		(0.003)		(0.003)
Farm heterogeneity	No	No	Yes	Yes	Yes	Yes
Regional fixed effects	No	No	No	No	District	District
Observations	216,118		216,118		216,118	
Hansen-J statistic	68.088***		6.100*		5.504	
Kleibergen-Paap rk LM statistic	431.151***		336.929***		259.889***	
Kleibergen-Paap rk Wald F statistic	633.983		362.911		186.783	

Notes: Standard errors are reported in parentheses and are clustered at the municipality level. *** significant at 0.1 %, ** significant at 1 %, * significant at 5 % level. Farm heterogeneity includes 24 variables on farm characteristics and 32 detailed farm type fixed effects. Parameter estimates for all farm characteristics are reported in Table B2 in Appendix B. The Hansen-J statistic is a test of overidentifying restrictions and tests if the instruments are valid instruments, i.e., uncorrelated with the error term. The Kleibergen-Paap rk LM statistic is an underidentification test to evaluate whether the equation is identified, i.e., whether the excluded instruments are correlated with the endogenous regressor *Agritourism*. The Kleibergen-Paap rk Wald F statistic tests whether the excluded instruments are only weakly correlated with the endogenous regressor *Agritourism*. The Hansen-J statistic and the Kleibergen-Paap rk LM statistic are distributed as χ^2 with one and two degrees of freedom, respectively. The Kleibergen-Paap rk Wald F statistics clearly exceed the critical values proposed by Stock and Yogo (2005).

can be expected. On the other hand, a (very) high number of beds may, in the long run, promote the exit from agriculture and the transition to a pure tourism business. In summary, these two countervailing effects warrant a more detailed analysis of the intensive margin.

We use information on the number of *Beds* for tourists for each farm (again from the FSS data) as an indicator to measure the intensity of agritourism on the farm. For the 15,458 farms offering tourist accommodation, the number of beds ranges from 1 to 179, with a mean of 11.0 and a standard deviation of 10.8 beds. The 5^{th} (95^{th}) percentile is 3 (30) beds. To determine the effects, we estimate a probit model on farm *Survival* for the sub-population of farms involved in *Agritourism*. All other variables are the same as in specification (3) of the main results Table 4.

Regression results for the coefficients from this analysis can be found in Table B3 in Appendix B.

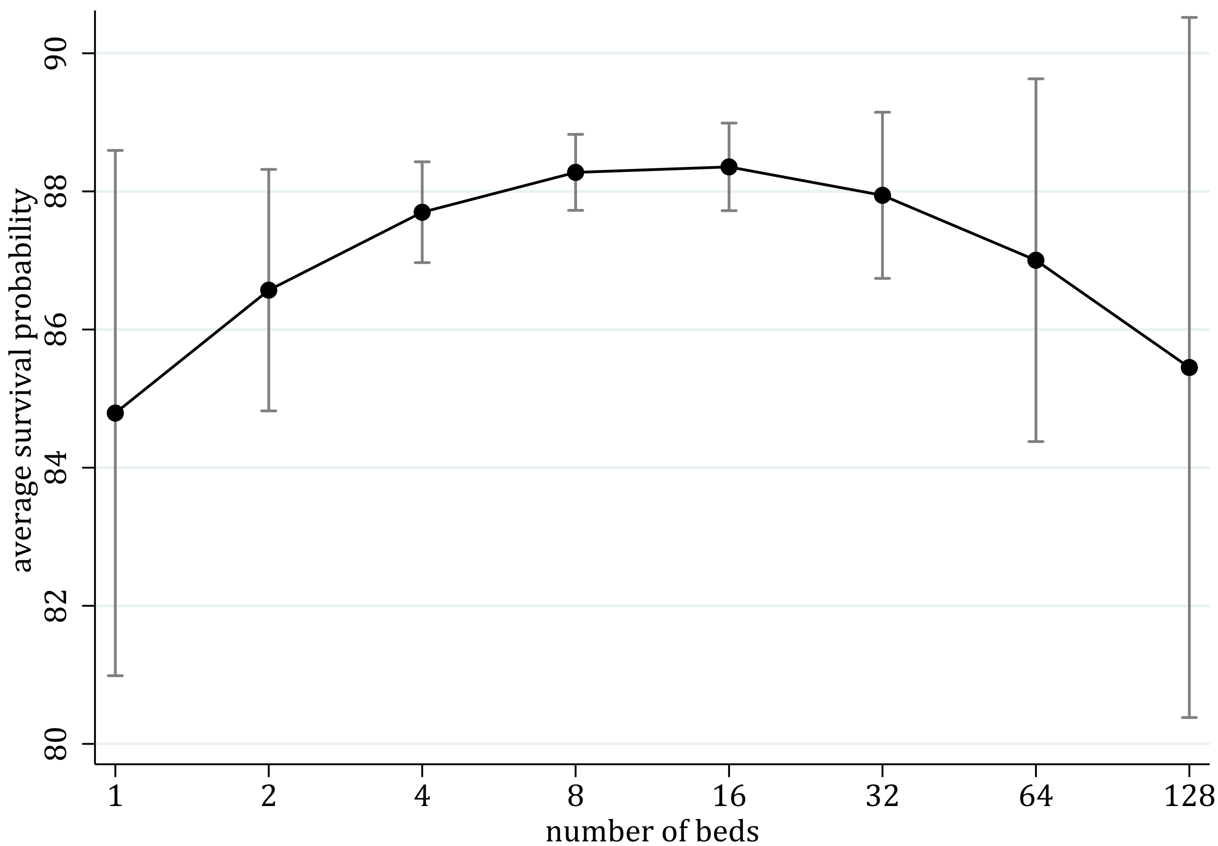
As reported in Table B3, the parameter estimate of $log(Beds)$ is positive but not significantly different from zero when this variable is included as a linear term only (specification (1)). When the variable $log(Beds)$ is considered in a linear-quadratic way (specification (2)), the coefficient of the linear term becomes significantly positive at the 10 % significance level, while the coefficient of $log(Beds)^2$ is negative (and is on the edge of being significantly different from zero at the 10 %-level). The point estimates suggest an inverse U-shaped relationship between the number of $Beds$ and farm $Survival$, albeit estimated with considerable statistical uncertainty.

The estimated survival probabilities depending on the number of beds are shown in Figure 4. Again, the results indicate that the effect between the number of beds and the $Survival$ probability is characterized by an inverse U and peaks at about 12.7 beds. The $Survival$ probabilities for farms with a very small or a rather large capacity are about 3 pp lower than for farms with about 10 to 15 beds. This is consistent with the expectation that farms offering only a small number of beds will have lower profitability and that many beds may encourage exit from agriculture in favor of a tourism-only business. However, the effects are not precisely estimated and do not significantly differ from each other according to the 95 % confidence interval for different numbers of beds, mainly due to the large confidence bands at the two ends of the distribution due to the small number of observations. In summary, although we find some evidence of heterogeneous effects, the extensive margin (offering agritourism) seems to be more important than the intensive margin (the number of beds a farm rents out).

5 Discussion and Conclusions

In this article, we investigate the impact of agritourism on farm survival. The empirical framework chosen allows a causal interpretation of the results and shows that offering farm-stay accommodation is an effective survival strategy. On average, agritourism increases the probability of survival by more than 10 percentage points over an eleven-year period and by more than 15 points for some sub-populations. It appears to be particularly effective for conventional farms in "normal" (not disadvantaged or difficult) locations, for part-time farms, and for farms managed by a farmer without formal agricultural training. Thus, entry into agritourism can be seen as a powerful means of

Figure 4: Predicted farm survival probabilities depending on the number of beds



Notes: The figure illustrates the predicted farm survival probabilities along with the 95 % confidence interval. Predictions are based on a probit estimation of farm survival for the sample of farms offering agritourism. Parameter estimates for all farm characteristics are reported in specification (2) of Table B3 in Appendix B. All variables except the number of beds are set to their sample means.

diversifying the farm portfolio, increasing the overall profitability of such farms beyond the threshold of market exit. Our analysis of the intensive margin suggests that the effect is not independent of the intensity of agritourism on the farm. The regression results indicate an inverse U-shaped relationship between the number of tourist beds and the probability of farm survival, with the largest positive effect occurring at a moderate size of about 12.7 tourist beds.

Our article also contributes to the discussion on how best to use the wealth of information (big data) that is becoming increasingly available in economic research in general and in tourism economics in particular. The empirical analysis is based on two waves of census data on Austrian farms. This is an extremely large data set covering the entire population of more than 200,000 Austrian farms and includes numerous farm and farmer characteristics. For each specification within the probit, recursive bivariate probit and linear probability models, the estimated effects of participation in agritourism on farm survival are highly significant and robust to the inclusion of additional farm characteristics or fixed effects at different regional levels once we control for the main farm and farm manager characteristics. However, the magnitude of the estimated coefficients is heavily biased as long as we do not adequately account for (endogenous) self-selection into agritourism: The simple probit model underestimates the effect of agritourism on farm survival by about 50 %.

It appears that unobserved farm or farmer characteristics that positively affect farm survival (e.g., the farmer's ability, effort, managerial or entrepreneurial skills) reduce the likelihood of offering agritourism. This explanation is intuitive, as the provision of farm-stay accommodation does not usually require special skills and abilities and can therefore provide an easily exploitable source of income for less productive farmers who would otherwise struggle to generate sufficient income from agricultural production alone. Moreover, this explanation is consistent with the negative correlation in the error terms when estimating of the determinants of participation in agritourism and farm survival in the recursive bivariate model, as indicated by the parameter $\rho < 0$ in Table 4. As pointed out by Mullainathan and Spiess (2017), using large data sets and adding more variables without an appropriate identification strategy can help to predict the outcome variables accurately, but it does not help to produce meaningful causal—and therefore policy-relevant—estimates of the key explanatory variables of interest.

Our findings have important policy implications, as agricultural subsidy programs for portfolio enlargement and vertical diversification aim to increase farm profitability to prevent farm exit. This article suggests that subsidizing entry into agritourism seems to be a promising tool for policy makers to help farms stay in the market, especially for conventional part-time farms that rely on on-farm income and farms that are managed by farmers without formal agricultural training. The effect appears to be greatest for farms with moderate tourism intensity, i.e., farms with limited capacity for tourists. Therefore, policies that promote farm survival by subsidizing investment in on-farm tourist accommodation should target "disadvantaged" farms (part-time farms, low formal education, etc.) and the creation of a moderate number of tourist beds, which also contributes to the preservation of the small-scale farming structure.

References

Bagi, F. S. and Reeder, R. J. (2012). Factors affecting farmer participation in agritourism. *Agricultural and Resource Economics Review*, 41(2):189 – 199.

Cameron, A. C. and Trivedi, P. K. (2005). *Microeconometrics: Methods and applications.* Cambridge university press.

Carpio, C. E., Wohlgenant, M. K., and Boonsaeng, T. (2008). The demand for agritourism in the United States. *Journal of Agricultural and Resource Economics*, 33(2):254 – 269.

Coban, M. (2020). Redistribution preferences, attitudes towards immigrants, and ethnic diversity. IAB-Discussion Paper 23/2020, Nürnberg.

Eurostat (2023). Farms and farmland in the European Union - statistics: Statistics explained. Technical report, Eurostat. https://ec.europa.eu/eurostat/statistics-explained/SEPDF/cache/73319.pdf, accessed Sept. 11 2023.

Evans, W. N. and Schwab, R. M. (1995). Finishing high school and starting college: Do catholic schools make a difference? *Quarterly Journal of Economics*, 110(4):941 – 974.

Goddard, E. W. (1993). Economics of structural change in agriculture. *Canadian Journal of Agricultural Economics*, 41(4):475 – 489.

Goetz, S. J. and Debertin, D. L. (2001). Why farmers quit: A county-level analysis. *American Journal of Agricultural Economics*, 83(4):1010 – 1023.

Greene, W. H. (2011). *Econometric Analysis.* Prentice Hall, Upper Saddle River, NJ, 7th edition.

Hallam, A. (1991). Economies of size and scale in agriculture: An interpretive review of empirical measurement. *Review of Agricultural Economics*, 13(1):155 – 172.

Hill, R., Loomis, J., Thilmany, D., and Sullins, M. (2014). Economic values of agritourism to visitors: A multi-destination hurdle travel cost model of demand. *Tourism Economics*, 20(5):1047 – 1065.

Jeczmyk, A., Uglis, J., Graja-Zwolinska, S., Mackowiak, M., Spychala, A., and Sikora, J. (2015). Research note: Economic benefits of agritourism development in Poland – an empirical study. *Tourism Economics*, 21(5):1120 – 1126.

Joo, H., Khanal, A. R., and Mishra, A. K. (2013). Farmers' participation in agritourism: Does it affect the bottom line? *Agricultural and Resource Economics Review*, 42(3):471 – 490.

Khanal, A. R. and Mishra, A. K. (2014). Agritourism and off-farm work: Survival strategies for small farms. *Agricultural Economics*, 45:65 – 76.

Kimhi, A. (2000). Is part-time farming really a step in the way out of agriculture? *American Journal of Agricultural Economics*, 82(1):38 – 48.

Kimhi, A. and Bollman, R. (1999). Family farm dynamics in Canada and Israel: The case of farm exits. *Agricultural Economics*, 21(1):69 – 79.

Lewbel, A., Dong, Y., and Yang, T. T. (2012). Comparing features of convenient estimators for binary choice models with endogenous regressors. *Canadian Journal of Economics*, 45(3):809 – 829.

Lupi, C., Giaccio, V., Mastronardi, L., Giannelli, A., and Scardera, A. (2017). Exploring the features of agritourism and its contribution to rural development in Italy. *Land use policy*, 64:383–390.

McGehee, N. G. and Kim, K. (2004). Motivation for agri-tourism entrepreneurship. *Journal of travel research*, 43(2):161–170.

Miller, M. K. (2021). The uses and abuses of matching in political science. George Washington University, Washington D.C.

Mullainathan, S. and Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, 31(2):87–106.

Nickerson, N. P., Black, R. J., and McCool, S. F. (2001). Agritourism: Motivations behind farm/ranch business diversification. *Journal of Travel research*, 40(1):19–26.

Ohe, Y. and Ciani, A. (2011). Evaluation of agritourism activity in Italy: Facility based or local culture based? *Tourism Economics*, 17(3):581 – 601.

Ohe, Y. and Ciani, A. (2012). Accessing demand characteristics of agritourism in Italy. *Tourism & Hospitality Management*, 18(2):281–296.

Pfeffer, M. J. (1989). Part-time farming and the stability of family farms in the Federal Republic of Germany. *European Review of Agricultural Economics*, 16(4):425 – 444.

Roe, B. (1995). A study of us farm exits with evidence from the panel study of income dynamics. Technical

report, Paper presented at the 1995 Annual meeting of the American Agricultural Economics Association.
Santeramo, F. G. (2015). Research note: Promoting the international demand for agritourism: Empirical evidence from a dynamic panel data model. *Tourism Economics*, 21(4):907 – 916.
Schilling, B. J., Attavanich, W., and Jin, Y. (2014). Does agritourism enhance farm profitability? *Journal of Agricultural and Resource Economics*, 39(1):69 – 87.
Sinabell, F., Bock-Schappelwein, J., Firgo, M., Friesenbichler, K. S., Piribauer, P., Streicher, G., Gerner, L., Kirchner, M., Kantelhardt, J., Niedermayr, A., Schmid, E., Schoenhart, M., and Mayer, C. (2019). *Eine Zwischenbilanz zu den Wirkungen des Programms der Ländlichen Entwicklung 2014-2020*. WIFO.
Stock, J. and Yogo, M. (2005). Testing for weak instruments in linear iv regression. In Andrews, D. W., editor, *Identification and Inference for Econometric Models*, pages 80–108. Cambridge University Press, New York.
Van der Ploeg, J. D., Roep, D., et al. (2003). Multifunctionality and rural development: The actual situation in europe. In van Huylenbroeck, G. and Durand, G., editors, *Multifunctional Agriculture; A new paradigm for European Agriculture and Rural Development*, chapter 3, pages 37–54. Ashgate Aldershot.
Weiss, C. R. (1996). Exits from a declining sector: Econometric evidence from a panel of Upper-Austrian farms 1980-90. Working Paper 9601, Department of Economics, University of Linz, Linz.
Weiss, C. R. (1999). Farm growth and survival: Econometric evidence for individual farms in Upper Austria. *American Journal of Agricultural Economics*, 81(1):103 – 116.
Wooldridge, J. M. (2001). *Econometric analysis of cross section and panel data.* Cambridge and London:.
Yeboah, A., Owens, J., Bynum, J., and Okafor, R. (2017). Factors influencing agritourism adoption by small farmers in North Carolina. *Journal of Agricultural Extension and Rural Development*, 9(5):84–96.
Zimmermann, A. and Heckelei, T. (2012). Structural change of European dairy farms–a cross-regional analysis. *Journal of Agricultural Economics*, 63(3):576 – 603.

A Details on farm-level control variables

A.1 Descriptive Statistics

Table A1 shows descriptive statistics on the FSS farm-level data that are included in the analysis as outcome and explanatory variables. This section describes and discusses these variables in detail.

A number of continuous and categorical variable enter both equations of the bivariate probit model (as well as the simple probit and the linear probability model) as control variables. *AlpinePasture* is a dummy variable equal to one if the land use of a farm includes alpine pastures, an is zero otherwise. 18.2 % of farms engaging in tourism have alpine pastures, while this figure is significantly smaller (4.8 %) not providing farm-stay accommodation. The Standard Gross Margin (*SGM*) is an indicator for determining the business orientation and economic size of farm activities. For each crop or animal production, *SGM* is calculated as the difference between the standard value of the production and the standard amount of certain specific costs (mainly proportional specific costs). The variable *SGM* is therefore an indicator of a farm's expected gross margins based on its inputs. The mean *SGM* is about 16,053 EUR in the total population, and is significantly lower by nearly 2,000 EUR among farms offering tourist accommodation. Two variables measure farm size in terms of labor as a production factor: *LaborForce* is the total labor force (headcount) and includes both family members and other employees who work on the farm in a professional capacity (full-time or part-time). *FamilyLF*, on the other hand, is the family labor force on the farm and includes available family members who work (help out) on the farm at least occasionally (whether professional or not). As shown in Table A1, the number of *FamilyLF* (3.8) is on average higher than that of *LaborForce* (2.6), but *FamilyLF* can be zero in many cases (e.g., for farms run as legal entities), while the minimum number of *LaborForce* is one. Both *LaborForce* and *FamilyLF* are significantly higher on average for farms engaged in tourism. *RedAgriArea* is the "Reduced Utilized Agricultural Area" in hectares with a mean of 11.9, and with agritourism farms slightly (but significantly) smaller than other farms.[20] *LeasedLandShare* is the share of farmland leased out in the farmland owned by the farm. The average share is 2.7 % among agritourism farms and

[20]The "Reduced Utilized Agricultural Area" includes land with normal yield potential (arable land, gardens, orchards, vineyards, vine and tree nurseries, forest tree nurseries, meadows mown several times, cultivated pastures) and extensively used permanent grassland areas (meadows mown once, rough grazings, litter meadows, alpine pastures and mountain meadows) converted by reduction factors.

Table A1: Descriptive statistics of farm-level variables

Variable	Agritourism Mean	Agritourism Std. Dev.	Agritourism Min	Agritourism Max	No Agritourism Mean	No Agritourism Std. Dev.	No Agritourism Min	No Agritourism Max	Total Mean	Total Std. Dev.	Total Min	Total Max	Diff.
Survivial	0.8803	0.3246	0.0	1.0	0.7364	0.4406	0.0	1.0	0.7467	0.4349	0.0	1.0	***
Agritourism	1.0000	0.0000	1.0	1.0	0.0000	0.0000	0.0	0.0	0.0715	0.2577	0.0	1.0	***
AlpinePasture	0.1817	0.3856	0.0	1.0	0.0480	0.2138	0.0	1.0	0.0576	0.2330	0.0	1.0	***
SGM (in 1,000 EUR)	14.2341	56.1461	0.0	5,370.7	16.1934	70.5828	0.1	10,910.0	16.0532	69.6514	0.0	10,910.0	***
LaborForce	3.0195	1.4708	1.0	53.0	2.6087	2.1272	1.0	240.0	2.6381	2.0898	1.0	240.0	***
FamilyLF	4.4375	1.8758	0.0	12.0	3.7828	2.0635	0.0	12.0	3.8297	2.0576	0.0	12.0	***
RedAgriArea	11.1584	13.3181	0.0	486.6	11.9560	22.8101	0.0	3,042.9	11.8989	22.2669	0.0	3,042.9	***
LeasedLandShare	0.0272	0.1065	0.0	1.0	0.0556	0.1646	0.0	1.2	0.0536	0.1613	0.0	1.2	***
Husbandry	0.8267	0.3785	0.0	1.0	0.6379	0.4806	0.0	1.0	0.6514	0.4765	0.0	1.0	***
LivestockU	13.6203	16.9914	0.0	604.4	14.4982	78.4512	0.0	12,412.5	14.4354	75.7303	0.0	12,412.5	***
Female	0.2165	0.4118	0.0	1.0	0.2909	0.4542	0.0	1.0	0.2856	0.4517	0.0	1.0	***
Age													
<50	0.5653	0.4957	0.0	1.0	0.5873	0.4923	0.0	1.0	0.5857	0.4926	0.0	1.0	***
50-59	0.2590	0.4381	0.0	1.0	0.2260	0.4182	0.0	1.0	0.2284	0.4198	0.0	1.0	***
60+	0.1758	0.3806	0.0	1.0	0.1867	0.3897	0.0	1.0	0.1859	0.3890	0.0	1.0	***
Training													
None	0.5508	0.4974	0.0	1.0	0.6282	0.4833	0.0	1.0	0.6227	0.4847	0.0	1.0	***
Basic	0.3384	0.4732	0.0	1.0	0.2739	0.4460	0.0	1.0	0.2785	0.4483	0.0	1.0	***
Higher	0.1108	0.3139	0.0	1.0	0.0979	0.2972	0.0	1.0	0.0988	0.2984	0.0	1.0	***
EnterpriseType													
Full-time	0.4537	0.4979	0.0	1.0	0.3623	0.4807	0.0	1.0	0.3688	0.4825	0.0	1.0	***
Part-time	0.5414	0.4983	0.0	1.0	0.5996	0.4900	0.0	1.0	0.5954	0.4908	0.0	1.0	***
Legal entity	0.0049	0.0695	0.0	1.0	0.0381	0.1915	0.0	1.0	0.0358	0.1857	0.0	1.0	***
DifficultyZone													
Level 0	0.2558	0.4363	0.0	1.0	0.6320	0.4823	0.0	1.0	0.6051	0.4888	0.0	1.0	***
Level 1	0.1756	0.3805	0.0	1.0	0.1194	0.3242	0.0	1.0	0.1234	0.3289	0.0	1.0	***
Level 2	0.2011	0.4008	0.0	1.0	0.1001	0.3002	0.0	1.0	0.1074	0.3096	0.0	1.0	***
Level 3	0.2771	0.4476	0.0	1.0	0.1237	0.3292	0.0	1.0	0.1346	0.3413	0.0	1.0	***
Level 4	0.0905	0.2869	0.0	1.0	0.0248	0.1555	0.0	1.0	0.0295	0.1692	0.0	1.0	***
DisadvArea													
None	0.0438	0.2046	0.0	1.0	0.3122	0.4634	0.0	1.0	0.2930	0.4551	0.0	1.0	***
Mountainous	0.8862	0.3176	0.0	1.0	0.4892	0.4999	0.0	1.0	0.5176	0.4997	0.0	1.0	***
Other disadv.	0.0388	0.1930	0.0	1.0	0.0901	0.2863	0.0	1.0	0.0864	0.2809	0.0	1.0	***
Small area	0.0312	0.1740	0.0	1.0	0.1085	0.3111	0.0	1.0	0.1030	0.3040	0.0	1.0	***
OEPUL													
Conventional	0.1457	0.3528	0.0	1.0	0.2956	0.4563	0.0	1.0	0.2849	0.4514	0.0	1.0	***
Organic	0.2892	0.4534	0.0	1.0	0.0717	0.2580	0.0	1.0	0.0873	0.2822	0.0	1.0	***
Sustainable non-organic	0.5651	0.4958	0.0	1.0	0.6326	0.4821	0.0	1.0	0.6278	0.4834	0.0	1.0	***

Note: *** Difference in mean between agritourism farms (15,458 obs.) and non-agritourism farms (200,660 obs.) significant at the 0.1 % level.

5.6 % among other farms, which is significantly higher. *Husbandry* is a dummy variable equal to one if a farm keeps livestock, and is zero otherwise. The former is the case for about 65.1 % of all farms. The ratio is substantially and statistically higher (82.7 %) among farms engaging in tourism. *LivestockU* is the number of standardized livestock units. By definition it is equal to zero if $Husbandry = 0$, and strictly larger than zero otherwise. For the total population of farms, the average *LivestockU* is equal to 14.4. Interestingly, the mean value is significantly lower for agritourism farms (13.6), although the share of husbandry is higher in this sub-population of farms. This implies that conditional on $Husbandry = 1$ farms not engaging in tourism have larger numbers of *LivestockU* on average than agritourism farms.

A number of categorical variables account for structural differences that are likely to impact both farm *Survival* and engaging in *Agritourism*. Of great importance—especially for survival—are characteristics of the type of business and the farm manager. *Female* is a dummy variable that has the value one if the farm is a family farm (i.e., not a legal entity) and the farm manager is female. This is the case in 21.7 % (29.1 %) of farms (not) participating in agritourism. *Age* of the manager is supposed to be a central determinant of survival probabilities (controlling for *FamilyLF*). We construct a categorical variable to account for the most likely differential effects of age on survival and participation in agritourism for different cohorts.[21] The reference category is legal entity farms and family-run farms with a manager younger than fifty. This applies to 58.6 % of all farms, while 22.8 % (18.6 %) have a manager with 50-59 (60+) years of age. Older cohorts are supposed to have a higher risk of exit if no family members are willing to take over the business upon retirement of the farm manager.[22] *Training* is a categorical variable indicating whether the farm manager has obtained a formal agricultural education degree. 55.1 % of farm managers in agritourism have exclusively practical agricultural training, but no formal technical education. The rate is substantially higher at 62.8 % among non-agritourism farms. In both groups of farms, the majority of those with formal agricultural education have received basic training (33.8 % and 27.4 %, respectively). Only in 11.1 % (9.8 %) of the agritourism (non-agritourism) farms did the farm managers receive a higher formal agricultural education. *EnterpriseType* provides information

[21] A total of 212 observations with missing information on the year of birth of the farm manager is excluded from the analysis.

[22] Note that the unique farm identifier that is used to determine survival and exit in the data does only change if a farm is taken over by another farm or legal entity, but not if a family member succeeds the previous farm manager (e.g., due to retirement).

whether the agricultural activity of the farm is carried out as a full-time farm business, as a part-time business or as a holding of a legal person.[23] Part-time business (59.5 %) is more common than full-term business (36.9 %). The relation is significantly different for agritourism farms (54.1 % part-time, 45.4 % full-time). Legal entities account only for 0.5 % of all farms engaged in tourism, compared to 3.6 % of the total farm population.

The variable $DifficultyZone$ is measured using a five-point Likert scale indicating the degree of difficulty of cultivation due to slopes at the farm level. It is an essential basis for the level of agricultural subsidies. A score of 0 (4) indicates low (high) difficulty. While in the total farm population 60.5 % score no difficulty, this is only the case for 25.6 % among farms offering agritourism. Together with the higher share of $AlpinePasture$, this shows that farms offering farm-stay accommodation are more often found in alpine regions. Similarly, $DisadvArea$ is a categorical variable for disadvantaged (less-favored) areas at the local (mainly municipality) level rather than the farm level. Only 4.4 % of agritourism farms are not located in disadvantaged areas, while this applies to 31.2 % of the remaining farms. Disadvantaged areas are divided into mountainous areas, small territories and other less-favored areas where the maintenance of a minimum population (density) or the preservation of the landscape would not be ensured without subsidies. $OEPUL$ is a shortcut for *Österreichisches Programm für umweltgerechte Landwirtschaft* (Austrian Program for Environmentally Friendly Agriculture). In this program, measure *OEPUL A* represents organic farming and measure *OEPUL B* represents other low-input farming systems or practices (integrated, ecological, etc.) that are considered environmentally sustainable. *OEPUL B* is therefore labeled as sustainable non-organic. Only about 14.6 % of farms in agritourism are considered conventional farms (i.e., neither *OEPUL A* nor *OEPUL B)*, while this is the case for 29.6 % of non-agritourism farms. Organic farms account for 28.9 % (7.2 %) of farms active (not active) in agritourism.

Table A1 shows substantial differences between farms that offer accommodation and those that do not, both in terms of farm and location characteristics. Therefore, including a large number of control variables as well as instrumenting the decision to provide touristic services is necessary to avoid omitted variable bias and self-selection bias. To further account for structural differences

[23] A full-time farm is a business where the farm manager couple worked more than 50 % of the total working time of the survey year on the agricultural and forestry holding and a minimum standard gross margin of 6,540 EUR was obtained. The three types are mutually exclusive. A part-time farm is defined as a farm where the farm manager couple worked less than 50 % of the total working time on the agricultural and forestry holding. Thus, off-farm employment accounted for at least 50 % of the total working time.

potentially impacting farm *Survival* and/or the decision to engage in *Agritourism*, we also include 32 dummy variables from a detailed typology of farms according to their agricultural production focus, dividing farms into sub-categories of field crop, dairy, permanent crop, mixed agricultural, horticultural, forestry farms, mixed farms and non-classifiable farms.[24] Descriptive statistics for these production focus fixed effects are not displayed in Table A1 for brevity.

A.2 Detailed discussion of the main results

Table A2 shows the parameter estimates for all variables (except farm type and regional fixed effects) for both regressions explaining farm *Survival* and participation in *Agritourism* under the recursive bivariate probit model. These results form the basis for calculating the marginal effects reported in Table 4. In this section, we focus on the influence of the explanatory variables, which are not discussed in the main part of the article.

A higher standard gross margin (*SGM*) is an indicator for the expected profitability of farming activities. Therefore, the significant positive coefficient of *SGM* for farm *Survival* is not particularly surprising. In addition, a higher *SGM* also increases the probability for participation in *Agritourism* once district fixed effects are included. Having an *AlpinePasture* significantly reduces the probability of *Survival* but does not significantly contribute to explaining participation in *Agritourism*. The influence on farm *Survival* is plausible because preservation of alpine pastures is typically laborious. It is somewhat surprising that the presence of *AlpinePastures* is not related to participation in *Agritourism*, so it does not seem to be necessarily important to provide tourists with their own alpine pastures, but to be located in an alpine environment (see also coefficient of *DisadvArea*).

Farm size plays an important role for both dependent variables: Both variables on farm employment are significantly positively related to farm *Survival* and *Agritourism* participation rates. Farms with higher labor force (*LaborForce*) and with more family workers (*FamilyLF*) are significantly more likely to survive and to engage in tourism. Higher *RedAgriArea* (Reduced Utilized Agricultural Area) increases the probability of farm *Survival*, but has little influence on the

[24] The typology originally considers 34 types. Two types in horticulture had to be combined because participation in *Agritourism* could be perfectly predicted due to the small number of observations in these categories. 280 farms classified as "not assignable" are excluded from the analysis, in part because failure to assign them to a category meant that the *SGM* could not be calculated for these farms in the data set.

Table A2: All parameter estimates of recursive bivariate probit models

	Model (1)		Model (2)		Model (3)	
Dependent variable	*Survival*	*Agri-tourism*	*Survival*	*Agri-tourism*	*Survival*	*Agri-tourism*
Agritourism	0.830***		0.395***		0.487***	
	(0.0720)		(0.0604)		(0.0598)	
LowTourism		−1.071***		−0.703***		−0.519***
		(0.0736)		(0.0708)		(0.0690)
ln(*OvernightStays*)		0.420***		0.433***		0.365***
		(0.00873)		(0.00912)		(0.0109)
ln(*BedsPerFacility*)		−0.433***		−0.355***		−0.275***
		(0.0268)		(0.0261)		(0.0245)
ln(*SGM*)			0.145***	0.0149*	0.144***	0.0233**
			(0.00437)	(0.00728)	(0.00431)	(0.00711)
AlpinePasture			−0.179***	0.0423	−0.213***	−0.00468
			(0.0287)	(0.0241)	(0.0262)	(0.0233)
ln(*LaborForce*)			0.0998***	0.107***	0.0952***	0.101***
			(0.00994)	(0.0159)	(0.00980)	(0.0157)
ln(*FamilyLF*)			0.272***	0.134***	0.273***	0.144***
			(0.00856)	(0.0157)	(0.00855)	(0.0153)
ln(*RedAgriArea*)			0.0530***	0.0173	0.0714***	0.0261*
			(0.00700)	(0.0107)	(0.00687)	(0.0102)
RedAgriArea Zero Flag			0.134***	−0.165***	0.171***	−0.153***
			(0.0172)	(0.0335)	(0.0170)	(0.0324)
LeasedLandShare			−0.0722**	0.199***	0.00241	0.240***
			(0.0249)	(0.0516)	(0.0230)	(0.0497)
Husbandry			−0.194***	−0.0232	−0.209***	−0.0632*
			(0.0156)	(0.0280)	(0.0153)	(0.0280)
ln(*LivestockU*)			0.178***	0.0646***	0.184***	0.0481***
			(0.00805)	(0.0119)	(0.00808)	(0.0114)
Age (reference: <50)						
50-59			−0.215***	0.118***	−0.218***	0.114***
			(0.00931)	(0.0125)	(0.00933)	(0.0125)
60+			−0.336***	0.0672***	−0.343***	0.0609***
			(0.00979)	(0.0153)	(0.00974)	(0.0153)
Female			−0.114***	0.0241	−0.110***	0.0355**
			(0.00773)	(0.0131)	(0.00765)	(0.0128)
Training (reference: none)						
Basic			0.0569***	0.105***	0.0454***	0.113***
			(0.00935)	(0.0142)	(0.00913)	(0.0136)
Higher			0.199***	0.212***	0.184***	0.198***
			(0.0162)	(0.0201)	(0.0161)	(0.0196)
EnterpriseType (reference: full-time)						
Part-time			−0.243***	−0.0653***	−0.240***	−0.0708***
			(0.0115)	(0.0151)	(0.0114)	(0.0150)
Legal entity			−0.0189	−1.162***	0.0241	−1.141***
			(0.0282)	(0.0715)	(0.0282)	(0.0705)

Table A2: All parameter estimates of recursive bivariate probit models (continued)

DifficultyZone (reference: Level 0)						
Level 1			−0.0192	0.0713*	0.0282	0.144***
			(0.0172)	(0.0281)	(0.0167)	(0.0259)
Level 2			0.137***	0.0284	0.173***	0.126***
			(0.0194)	(0.0285)	(0.0177)	(0.0266)
Level 3			0.246***	−0.0811**	0.273***	0.0260
			(0.0186)	(0.0288)	(0.0177)	(0.0271)
Level 4			0.487***	−0.225***	0.482***	−0.156***
			(0.0296)	(0.0385)	(0.0294)	(0.0376)
DisadvArea (reference: none)						
Mountainous			0.236***	0.297***	0.132***	0.240***
			(0.0185)	(0.0437)	(0.0220)	(0.0503)
Other disadv.			0.0812**	0.0783	0.0821**	0.117*
			(0.0270)	(0.0481)	(0.0255)	(0.0557)
Small area			0.306***	0.101	0.122***	0.103
			(0.0226)	(0.0528)	(0.0233)	(0.0576)
OEPUL (reference: conventional)						
Organic (OEPUL A)			0.650***	0.394***	0.663***	0.334***
			(0.0224)	(0.0270)	(0.0219)	(0.0250)
Sustainable non-organic (OEPUL B)			0.370***	0.120***	0.386***	0.101***
			(0.0126)	(0.0214)	(0.0124)	(0.0203)
farm type fixed effects	No	No	Yes	Yes	Yes	Yes
Regional fixed effects	No	No	No	No	District	District
Observations	216,118		216,118		216,118	
log-likelihood	−163,856.9		−132,105.6		−130,473.3	

Notes: Results estimated using the Stata command *rbiprobit* (Coban, 2020).
Standard errors are reported in parentheses and are clustered at the municipality level. *** significant at 0.1 %, ** significant at 1 %, * significant at 5 % level.

decision to offer farm-stay accommodation. Farms with $RedAgriArea = 0$—indicated by the corresponding zero flag dummy variable—have a significantly higher (lower) probability for *Survival* (*Agritourism*).

Leasing out more land (*LeasedLandShare*) is not associated with higher *Survival* probabilities. The coefficient for this indicator is insignificant in specification (3) and even significantly negative in specification (2). However, a higher share of land leased out leads increases the probability for participation in *Agritourism*, implying that there is a substitution between land cultivation and provision of tourism accommodation. Farms with *Husbandry* have a lower probability for *Survival* but not for participation in *Agritourism*, ceteris paribus, but among *Husbandry* farms both *Survival* and *Agritourism* probabilities increase with higher numbers of livestock (*LivestockU*).

Characteristics of the farm manager and legal forms also play a significant role in both equations of the bivariate probit model. With *Age* <50 as a reference group, *Survival* probabilities are significantly lower in farm manager age cohorts 50-59 and 60+, with the oldest cohort having the lowest *Survival* probabilities. Similarly, farms with a *Female* manager have significantly lower *Survival* probabilities (compared to farms with a male manager serving as a reference group). While gender is not associated with a strong impact on participation probabilities in *Agirtourism*, age groups above 50 both are significantly positive related to the probability to engage in tourism compared to younger age cohorts. Formal *Training* in agriculture increases probabilities of both farm *Survival* and participation in *Agritourism* compared to the reference category of only practical on-farm training. In addition, higher formal education in agriculture increases both probabilities to a higher extent than basic formal agricultural education. Compared to farms with a farmer working full-time at the farm, part-time farms have both lower *Survival* and *Agritourism* participation probabilities, as indicated by *EnterpriseType*. Farms that constitute legal entities do not have other *Survival* probabilities than full-time farms, ceteris paribus, but have significantly lower probabilities to engage in *Agritourism*.

Farms at locations with natural disadvantages (*DifficultyZone*) and disadvantaged geographic areas (*Disadv Area*) have higher survival rates, ceteris paribus, than farms in locations without land in cumbersome zones and in non-disadvantaged geographic areas, respectively. This is not surprising given both limited outside options in very peripheral locations and the rather high amount of public subsidies dedicated to Austrian farms in zones difficult to cultivate and in disadvantaged areas (Sinabell et al., 2019). While *Survival* probabilities increase with the level of *DifficultyZone*, ceteris paribus, participation rates in *Agritourism* do not. According to the preferred specification (3), only farms in moderately difficult zones (level 1 and 2) have higher participation probabilities than farms in non-difficulty zones (the reference group). Farms in level 3 zones do not significantly differ from the latter. Farms in zones with the highest degree of difficulty (level 4) eventually have lower probabilities for participation in *Agritourism*. This may indicate that farms in level 4 zones might be located in sites that are difficult to access for potential tourists by means of transportation and/or too far from complementary tourism infrastructure. Among farms in disadvantaged geographic areas (*DisadvArea*), particularly those in mountainous areas have a higher ceteris paribus probability for participation in *Agritourism*, but also those in small and other disadvantaged areas have higher probabilities compared to the reference group not located in disadvantaged areas.

Organic and sustainable non-organic farms have significantly higher *Survival* probabilities and are more likely to engage in *Agritourism*. Sustainable farms (and in particular organic farms) have substantially higher probabilities both to survive and to provide tourism accommodation. Sustainable farms (and especially organic farms) are much more likely to both survive and provide tourist accommodations than conventional farms.

With respect to the impact of tourism activities at the municipality level for the individual farm-level decision to offer *Agritourism*, all variables have the expected sign. The coefficient of the identifying dummy variable *LowTourism* is significant and negative: Thus, farms in low-tourism municipalities have a significantly lower probability to engage in tourism, ceteris paribus. The number of overnight stays (*MunTourismStays*) in the municipality increases the farm-level probability for *Agritourism*. The average size of tourism establishments (*MunEstablSize*), measured by the average number of beds per accommodation, is significantly negative: The larger the average accommodation in the municipality, the lower the probability for farms to engage in *Agritourism*. This means that the composition effect (higher demand for agritourism in municipalities with small-scale tourism facilities) outweighs the competition effect (small-scale facilities are better substitutes for agritourism than large-scale hotels and resorts).

B Appendix Tables

Table B1: All parameter estimates of simple probit models

Dependent variable	Model (1) *Survival*	Model (2) *Survival*	Model (3) *Survival*	Model (4) *Survival*
Agritourism	0.544***	0.177***	0.179***	0.177***
	(0.0194)	(0.0185)	(0.0177)	(0.0189)
ln(*SGM*)		0.145***	0.145***	0.152***
		(0.00439)	(0.00434)	(0.00437)
AlpinePasture		−0.159***	−0.204***	−0.205***
		(0.0279)	(0.0258)	(0.0272)
ln(*LaborForce*)		0.102***	0.0991***	0.0964***
		(0.00994)	(0.00980)	(0.0100)
ln(*FamilyLF*)		0.274***	0.277***	0.290***
		(0.00857)	(0.00852)	(0.00870)
ln(*RedAgriArea*)		0.0521***	0.0707***	0.0737***
		(0.00706)	(0.00692)	(0.00709)
RedAgriArea Zero Flag		0.135***	0.171***	0.187***
		(0.0173)	(0.0170)	(0.0175)
LeasedLandShare		−0.0724**	0.00557	0.0429
		(0.0250)	(0.0230)	(0.0237)
Husbandry		−0.194***	−0.211***	−0.223***
		(0.0157)	(0.0154)	(0.0157)
ln(*LivestockU*)		0.180***	0.187***	0.200***
		(0.00810)	(0.00812)	(0.00837)
Age (reference: <50)				
50-59		−0.212***	−0.216***	−0.222***
		(0.00933)	(0.00935)	(0.00958)
60+		−0.335***	−0.342***	−0.353***
		(0.00978)	(0.00973)	(0.00997)
Female		−0.116***	−0.111***	−0.106***
		(0.00772)	(0.00766)	(0.00787)
Training (reference: none)				
Basic		0.0581***	0.0484***	0.0513***
		(0.00934)	(0.00912)	(0.00931)
Higher		0.203***	0.190***	0.197***
		(0.0162)	(0.0161)	(0.0164)
EnterpriseType (reference: full-time)				
Part-time		−0.245***	−0.243***	−0.245***
		(0.0115)	(0.0114)	(0.0117)
Legal entity		−0.0334	−0.00225	0.0266
		(0.0280)	(0.0282)	(0.0289)
DifficultyZone (reference: none)				
Level 1		−0.0181	0.0383*	0.0458*
		(0.0173)	(0.0167)	(0.0184)
Level 2		0.141***	0.186***	0.186***
		(0.0196)	(0.0179)	(0.0195)
Level 3		0.251***	0.286***	0.290***
		(0.0186)	(0.0175)	(0.0197)
Level 4		0.501***	0.488***	0.511***
		(0.0294)	(0.0293)	(0.0327)
DisadvArea (reference: none)				
Mountainous		0.252***	0.139***	0.252***
		(0.0185)	(0.0222)	(0.0642)
Other disadv.		0.0848**	0.0851***	0.158*
		(0.0269)	(0.0255)	(0.0669)
Small area		0.309***	0.125***	0.150
		(0.0228)	(0.0234)	(0.132)
OEPUL (reference: conventional)				
Organic (OEPUL A)		0.678***	0.692***	0.718***
		(0.0215)	(0.0211)	(0.0221)
Sustainable non-organic (OEPUL B)		0.373***	0.391***	0.403***
		(0.0126)	(0.0125)	(0.0127)
Farm type fixed effects	No	Yes	Yes	Yes
Regional fixed effects	No	No	District	Municipality
Observations	216,118	216,118	216,118	216,053
log-likelihood	−121,400.5	−93,151.5	−92,165.7	−90,017.9

Notes: Standard errors are reported in parentheses and are clustered at the municipality level. *** significant at 0.1 %, ** significant at 1 %, * significant at 5 % level.

Table B2: All parameter estimates of linear probability models

Dependent variable	Model (1) *Survival*	Model (1) *Agri-tourism*	Model (2) *Survival*	Model (2) *Agri-tourism*	Model (3) *Survival*	Model (3) *Agri-tourism*
Agritourism	0.275***		0.102***		0.0894***	
	(0.0274)		(0.0189)		(0.0253)	
LowTourism		−0.0473***		−0.0269**		−0.0182*
		(0.00904)		(0.00859)		(0.00753)
ln(*OvernightStays*)		0.0725***		0.0699***		0.0556***
		(0.00212)		(0.00227)		(0.00250)
ln(*BedsPerFacility*)		−0.0439***		−0.0363***		−0.0279***
		(0.00331)		(0.00316)		(0.00269)
ln(*SGM*)			0.0540***	0.00154*	0.0540***	0.00186**
			(0.00126)	(0.000672)	(0.00125)	(0.000645)
AlpinePasture			−0.0642***	0.0346***	−0.0745***	0.0201***
			(0.00562)	(0.00648)	(0.00508)	(0.00598)
ln(*LaborForce*)			0.0239***	0.0135***	0.0221***	0.0134***
			(0.00243)	(0.00170)	(0.00236)	(0.00162)
ln(*FamilyLF*)			0.0710***	0.0112***	0.0714***	0.0126***
			(0.00224)	(0.00160)	(0.00223)	(0.00153)
ln(*RedAgriArea*)			0.0167***	−0.00519***	0.0218***	−0.00462***
			(0.00217)	(0.00116)	(0.00216)	(0.00109)
RedAgriArea Zero Flag			0.0409***	−0.00116	0.0498***	−0.000700
			(0.00573)	(0.00290)	(0.00562)	(0.00277)
LeasedLandShare			−0.0394***	0.0222***	−0.0169*	0.0212***
			(0.00838)	(0.00384)	(0.00760)	(0.00353)
Husbandry			−0.0232***	−0.00453*	−0.0290***	−0.00595**
			(0.00474)	(0.00202)	(0.00461)	(0.00191)
ln(*LivestockU*)			0.0120***	0.00851***	0.0145***	0.00779***
			(0.00190)	(0.00122)	(0.00191)	(0.00111)
Age (reference: <50)						
50-59			−0.0500***	0.0135***	−0.0501***	0.0119***
			(0.00234)	(0.00142)	(0.00232)	(0.00141)
60+			−0.104***	0.00830***	−0.105***	0.00672***
			(0.00299)	(0.00162)	(0.00294)	(0.00159)
Female			−0.0271***	−0.00153	−0.0250***	0.000386
			(0.00207)	(0.00126)	(0.00201)	(0.00117)
Training (reference: none)						
Basic			0.00897***	0.0106***	0.00536**	0.0117***
			(0.00212)	(0.00185)	(0.00206)	(0.00171)
Higher			0.0246***	0.0212***	0.0198***	0.0207***
			(0.00312)	(0.00237)	(0.00309)	(0.00226)
EnterpriseType (reference: full-time)						
Part-time			−0.0365***	−0.00945***	−0.0359***	−0.0100***
			(0.00254)	(0.00184)	(0.00250)	(0.00178)
Legal entity			0.0214*	−0.0947***	0.0277**	−0.0962***
			(0.00849)	(0.00557)	(0.00864)	(0.00565)
DifficultyZone (reference: none)						
Level 1			−0.00679	0.0239***	0.00627	0.0353***
			(0.00438)	(0.00392)	(0.00426)	(0.00372)
Level 2			0.0250***	0.0244***	0.0358***	0.0361***
			(0.00478)	(0.00455)	(0.00434)	(0.00429)
Level 3			0.0451***	0.00959*	0.0547***	0.0207***
			(0.00467)	(0.00465)	(0.00431)	(0.00436)
Level 4			0.105***	−0.0123	0.0987***	−0.0149
			(0.00630)	(0.00883)	(0.00595)	(0.00885)
DisadvArea (reference: none)						
Mountainous			0.0616***	−0.0101*	0.0336***	−0.0153**
			(0.00527)	(0.00414)	(0.00584)	(0.00474)
Other disadv.			0.0160*	−0.0154***	0.0235***	−0.0150**
			(0.00775)	(0.00467)	(0.00701)	(0.00536)
Small area			0.0822***	−0.00724	0.0361***	−0.00448
			(0.00617)	(0.00414)	(0.00631)	(0.00453)
OEPUL (reference: none)						
Organic (OEPUL A)			0.154***	0.0916***	0.159***	0.0777***
			(0.00512)	(0.00487)	(0.00489)	(0.00433)
Sustainable non-organic (OEPUL B)			0.130***	0.0113***	0.133***	0.0115***
			(0.00401)	(0.00182)	(0.00390)	(0.00170)
Farm type fixed effects	No	No	Yes	Yes	Yes	Yes
Regional fixed effects	No	No	No	No	District	District
Observations	216,118		216,118		216,118	
R^2	0.00123	0.145	0.247	0.182	0.255	0.195

Notes: Standard errors are reported in parentheses and are clustered at the municipality level.*** significant at 0.1 %, ** significant at 1 %, * significant at 5 % level.

Table B3: All parameter estimates of probit models for the intensive margin of agritourism

	Model (1) *Survival*	Model (2) *Survival*
ln(*Beds*)	0.0154	0.177
	(0.0252)	(0.0992)
ln(*Beds*)^2		−0.0348
		(0.0215)
ln(*SGM*)	0.137***	0.138***
	(0.0180)	(0.0180)
AlpinePasture	−0.00937	−0.00976
	(0.0502)	(0.0500)
ln(*LaborForce*)	0.116*	0.112*
	(0.0458)	(0.0460)
ln(*FamilyLF*)	0.236***	0.237***
	(0.0408)	(0.0408)
ln(*RedAgriArea*)	0.0446	0.0446
	(0.0291)	(0.0290)
RedAgriArea Zero Flag	0.370***	0.370***
	(0.0812)	(0.0813)
LeasedLandShare	−0.0967	−0.0999
	(0.126)	(0.126)
Husbandry	−0.329***	−0.330***
	(0.0789)	(0.0790)
ln(*LivestockU*)	0.259***	0.260***
	(0.0396)	(0.0394)
Age (reference: <50)		
50-59	−0.115**	−0.114**
	(0.0368)	(0.0369)
60+	−0.319***	−0.320***
	(0.0415)	(0.0414)
Female	−0.0792*	−0.0800*
	(0.0372)	(0.0372)
Training (reference: none)		
Basic	0.0669	0.0661
	(0.0392)	(0.0392)
Higher	0.192**	0.190**
	(0.0713)	(0.0714)
EnterpriseType (reference: full-time)		
Part-time	−0.162***	−0.158***
	(0.0447)	(0.0446)
Legal entity	−0.346	−0.323
	(0.201)	(0.202)
DifficultyZone (reference: none)		
Level 1	−0.00929	−0.0109
	(0.0612)	(0.0611)
Level 2	0.178**	0.177**
	(0.0612)	(0.0611)
Level 3	0.205***	0.204***
	(0.0617)	(0.0617)
Level 4	0.431***	0.432***
	(0.0798)	(0.0798)
DisadvArea (reference: none)		
Mountainous	0.346**	0.344**
	(0.106)	(0.106)
Other disadv.	0.0662	0.0682
	(0.126)	(0.126)
Small area	0.213	0.211
	(0.120)	(0.120)
OEPUL (reference: none)		
Organic (OEPUL A)	0.653***	0.651***
	(0.0659)	(0.0659)
Sustainable non-organic (OEPUL B)	0.572***	0.572***
	(0.0521)	(0.0520)
Farm type fixed effects	Yes	Yes
Regional fixed effects	District	District
Observations	15,419	15,419
log-likelihood	−4,186.3	−4,184.6

Notes: Standard errors are reported in parentheses and are clustered at the municipality level. *** significant at 0.1 %, ** significant at 1 %, * significant at 5 % level.