sustainability

MDPI

Article

Financing Women's Farming Entrepreneurship: The Role of Agricultural Lending, Land Access, and Regional Heterogeneity in Kazakhstan

Talantty Arzayeva, Baglan Aimurzina *, Zhanat Bulakbay, Gaukhar Mataibayeva, Ramazan Sadvakassov and Gaukhar Kodasheva

Department of Finance, Higher School of Economics and Tourism, L.N. Gumilyov Eurasian National University, 11 Kazhymukan Street, Astana 010000, Kazakhstan; talanttyarzayeva@gmail.com (T.A.); bulakbay_zhannat@mail.ru (Z.B.); mataibaevagauhar@mail.ru (G.M.); sadvakasov.rm@gmail.com (R.S.); kodasheva_gs@enu.kz (G.K.)
* Correspondence: aimurzina_bt@enu.kz

Abstract

Unequal access to agricultural resources remains a major obstacle to women's entrepreneurship and inclusive rural development in Central Asia, yet empirical evidence on how agricultural lending relates to women's farming entrepreneurship is limited. This study examines the relationship between institutional agricultural lending and the number of women-led farms in 20 regions of Kazakhstan in 2015–2024. Pooled ordinary least squares, two-way fixed effects, the between estimator and the Mundlak regression are used to separate differences between regions from changes within regions over time; k-means clustering and a Random Forest model are supplementary methods. The pooled model shows a positive and statistically significant association between lending and women-led farms. With region and year fixed effects, the estimated association becomes smaller and is not statistically significant, although a within-region relationship cannot be ruled out. The official indicator of women granted agricultural land varies widely across regions. Cluster analysis identifies two groups of regions that differ in the level and growth of women-led farms, while the Random Forest model shows limited predictive ability. The study provides regional panel evidence from Kazakhstan and highlights the importance of regional context, while incomplete metadata limit conclusions about women's access to land.

Keywords: agricultural lending; women's farming entrepreneurship; financial inclusion; regional heterogeneity; Kazakhstan; inclusive development; access to land; Sustainable Development Goals

Check for updates

Academic Editors: Karmen Pažek and Črtomir Rozman

Received: 20 September 2026
Revised: 25 October 2026
Accepted: 27 October 2026
Published: 29 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Expanding women's economic opportunities and rights in agriculture is seen as an important condition for improving well-being, productivity, and food security [1–3]. As agricultural systems are transformed, women's equal access to resources (land, finance, technology) becomes not only a matter of social justice but also a condition for increasing the productivity of the sector.

This issue is directly linked to the United Nations Sustainable Development Goals (SDGs): SDG 1 (ending poverty), SDG 2 (ending hunger) and SDG 5 (gender equality).

In the economic literature, increased access to financial resources is considered a key factor in the growth of entrepreneurship. Studies in Africa and South Asia show that

improving women's access to financial resources, microfinance, and credit mechanisms may be associated with higher entrepreneurial activity and improved agricultural performance [4–8]. However, the nature of this relationship remains a subject of intense debate.

The key problems of gender inequality in agriculture include unequal access to land [9,10], limited participation of women in decision-making [11–16], and social and institutional barriers [17,18].

Women face systemic constraints in their access to resources, markets, financing, and institutional support [19–24]. Women's microenterprises in particular often have limited access to financial and social capital, markets, and development opportunities [22,23].

However, the findings of studies conducted in African and South Asian countries [8,25–29] are not always applicable to post-Soviet countries such as Kazakhstan. According to the Bureau of National Statistics of Kazakhstan, the share of women granted agricultural land decreased from 4.06% in 2015 to 1.53% in 2024. In 2024, it was 0.72% in urban areas and 3.12% in rural areas [30]. These figures show marked variation in the published administrative indicator, but incomplete national metadata limit their interpretation as differences in women's access to land (Section 4.1).

A literature review reveals three interrelated limitations of existing research. First, the predominant use of cross-sectional data makes it impossible to fully distinguish between interregional differences and changes within regions over time, whereas panel methods can account for both the temporal and the individual heterogeneity of observations [31–33]. Second, research focuses mainly on countries in Africa and South Asia [8,25–29]. The post-Soviet and Central Asian contexts are much less represented. Third, studies of financing rarely address access to land, although in agriculture land is not only a productive resource but also a key collateral asset [2].

Thus, the literature shows that access to finance and access to agricultural land are important conditions for women's economic empowerment. However, much of the research examines these factors at the individual farm level or through indicators of productivity, income, and food security. It remains poorly understood whether the statistical relationship between agricultural lending and the development of women's farming entrepreneurship holds after accounting for persistent regional differences and common time effects. This question is particularly relevant for Kazakhstan, where the regions differ considerably in their economic structure, level of agricultural development, and institutional conditions. This study addresses this gap using regional panel data for Kazakhstan for 2015–2024.

Pooled ordinary least squares (OLS) and the two-way fixed effects (TWFE) model are used to compare pooled associations with within-region relationships after accounting for region and year fixed effects.

Cluster analysis and Random Forest are used as complementary tools to characterize regional heterogeneity and identify potential nonlinear relationships. These methods are not used to establish causal relationships.

This study aims to assess the relationship between agricultural lending, women's access to agricultural land, and the development of women's farming entrepreneurship in Kazakhstan. The main research questions are as follows:

1. Is there a statistically significant relationship between agricultural lending and the number of women-led farms in the regions of Kazakhstan?
2. Does this relationship persist after accounting for region and year fixed effects?
3. Are changes in women's access to agricultural land related to changes in the number of women-led farms?
4. Are there stable groups of regions that differ in the level and dynamics of women's farming entrepreneurship?

This study contributes to the literature in three interrelated ways:

1. It expands empirical research on women's farming by analyzing regional panel data for Kazakhstan for 2015–2024.
2. It distinguishes between interregional differences and intraregional dynamics by comparing Pooled OLS and TWFE estimates and by separating the two sources of variation with a between estimator and a Mundlak regression.
3. It complements the econometric analysis with k-means cluster analysis, which identifies groups of regions with similar characteristics, and Random Forest, which assesses the predictive importance of factors and potential nonlinear relationships.

The rest of the article is organized as follows. Section 2 presents the theoretical background, and Section 3 develops the hypotheses. Section 4 describes the research methodology (data sources, specification of the econometric models, cluster analysis, and Random Forest). Section 5 presents the results, Section 6 discusses them, and Section 7 presents the conclusions.

2. Theoretical Background

Proponents of the financial approach [34–36] regard expanding access to microfinance and banking services as one of the key mechanisms of women's empowerment. Wider economic opportunities are assumed to support women's entrepreneurial activity, increase their participation in decision-making and improve socio-economic outcomes [37–39].

The institutional approach [2,40–43], by contrast, emphasizes that the effectiveness of financial instruments may be limited without addressing land rights, market access, and women's participation in decision-making. Credit alone may be insufficient to support sustainable growth when institutional conditions are unfavorable. The institutional environment and the way producers are integrated into agri-food chains can influence their ability to create and capture economic value [44].

According to a number of studies, women's access to financial resources is shaped by broader social, resource and institutional constraints [45–48]. In addition, a lack of entrepreneurial skills and competencies can limit women's entrepreneurial activity and their ability to use the available economic resources effectively [49,50]. The importance of structural and institutional barriers to the development of women's entrepreneurship and financial inclusion is also examined in studies on Kazakhstan and Central Asia [51–54].

Empirical research conducted mainly in Africa and South Asia identifies three main mechanisms through which land rights affect women's entrepreneurship: (1) land assets can be used as collateral and thus expand access to formal loans [2]; (2) secure property rights encourage long-term investment in land improvement and the adoption of new technologies [2,55]; and (3) land ownership increases women's confidence in the future, reducing risks and contributing to the sustainable development of farms [2,10].

The financial and institutional approaches provide a theoretical framework for examining the relationship between agricultural credit, women's access to agricultural land, and women's entrepreneurship in agriculture. The financial approach emphasizes the potential role of credit in supporting entrepreneurship, while the institutional approach focuses on the conditions under which financial resources can be accessed and used. These perspectives suggest that the relationship between credit and women's entrepreneurship in agriculture should be examined alongside access to land and regional institutional differences.

The combination of these approaches provides a conceptual justification for including agricultural credit and women's access to agricultural land as explanatory variables. Regional heterogeneity is considered an important contextual aspect of the analysis.

The financial and institutional dimensions of women's farming entrepreneurship are also relevant to sustainable rural development and the resilience of farming systems. Farming system resilience encompasses the capacities to withstand disturbances, adapt to

changing conditions, and transform when necessary [56]. Innovation and technology can contribute to resilient and inclusive rural transformation [57]. These perspectives provide a broader sustainability context for examining agricultural lending, women's access to agricultural land, and regional heterogeneity.

3. Hypotheses Development

Previous studies suggest that access to microfinance and banking services may support women's entrepreneurial activity [8,34–36]. However, evidence from agricultural microcredit programs indicates that access to finance does not necessarily translate into improved agricultural outcomes [5]. These findings suggest that the relationship between agricultural lending and women's farming entrepreneurship requires further empirical examination. Accordingly, the first hypothesis is proposed:

H1. *Higher levels of agricultural lending are positively associated with the number of women-led farms.*

Women's land rights and access to agricultural resources are important for their economic opportunities [2]. Research also identifies gender differences in the valuation of agricultural land [9] and in the distribution and documentation of land rights [10]. These findings provide a basis for examining whether women's access to agricultural land is associated with the development of women's farming entrepreneurship. On this basis, the second hypothesis is formulated:

H2. *Greater access to agricultural land among women is positively associated with the number of women-led farms.*

Studies on women's entrepreneurship and financial inclusion in Kazakhstan point to the importance of structural and institutional barriers [52,54]. Research on women's economic participation in Central Asia also points to differences in economic conditions across countries [53]. Although these studies do not directly establish differences among Kazakhstan's regions, variations in local economic conditions and access to productive resources provide a rationale for examining whether women's farming entrepreneurship follows different regional patterns. This leads to the third hypothesis:

H3. *The regions of Kazakhstan can be grouped into clusters that differ in the level and dynamics of women's farming entrepreneurship, as well as in the characteristics of agricultural lending and women's access to agricultural land.*

4. Research Methodology

The study assesses the relationship between agricultural lending, institutional factors and the development of women's farming entrepreneurship. A quantitative approach was chosen to assess the statistical significance of the relationships, the differences between regions, and the dynamics of the indicators within regions. In addition, regions with similar characteristics are identified, and the predictive importance of the factors under consideration is evaluated.

4.1. Data and Variables

The analysis is based on an unbalanced panel of data for 20 regions of Kazakhstan for 2015–2024. The sample comprises 175 observations. It includes 17 regions and 3 cities of national significance (Astana, Almaty, and Shymkent), which are counted as separate regions in official statistics.

The period 2015–2024 was chosen based on the availability of comparable agricultural indicators. In 2022, three new regions were formed (Abai, Zhetysu and Ulytau); therefore, no data are available for them for earlier years.

The establishment of Abai, Zhetysu, and Ulytau regions in 2022 also changed the territorial boundaries of East Kazakhstan, Almaty, and Karaganda regions, respectively. These changes may affect the comparability of regional statistics before and after 2022.

The data are official statistics of the Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan.

The dependent variable is the number of peasant farms headed by women (women_farms), hereinafter referred to as women-led farms. The indicator is the annual number of such registered farms. It reflects only the number of farms and does not take into account their size, productivity, or profitability.

The main financial variable is the total amount of loans issued to women entrepreneurs in agriculture (credit_total), in millions of tenge. The indicator includes the volume of loans issued through three joint-stock companies (JSC): "KazAgroFinance" JSC (credit_kaf), "Agrarian Credit Corporation" JSC (credit_acc) and the "Fund for Financial Support of Agriculture" JSC (credit_fund). To check the robustness of the results, we also use these components separately.

The institutional variable (land_share) is the regional value of Kazakhstan's national SDG indicator 5.a.1, "Proportion of women granted agricultural land, by form of ownership" (%), for the total population of the region, as published by the Bureau of National Statistics [30]. The Bureau classifies it as a modified version of the global indicator 5.a.1 and compiles it from administrative data of the Land Management Committee of the Ministry of Agriculture and local executive bodies; the published series covers the total, urban and rural population, without a breakdown by form of ownership. The accompanying metadata reproduce the FAO metadata for the global indicator, which is designed for household surveys and whose sub-indicator 5.a.1(b) is the share of women among all agricultural land rights-holders, but they do not specify the numerator, the denominator or the reference period of the national indicator. The available documentation therefore does not establish the national calculation formula or its equivalence with SDG 5.a.1(b). We use the published values without recalculation and interpret land_share only as the published administrative indicator, not as women's share among land rights-holders or as the number of women granted land in a given year.

To account for the right skewness of the distributions and to handle zero values correctly, the dependent variable and all credit variables (prefixed log_ in the models) were transformed using the natural logarithm $\ln(1 + x)$. This transformation reduces the influence of extremely large values and retains observations with zero values.

The panel is unbalanced for two reasons: the administrative changes of 2022 and incomplete reporting of some indicators in certain regions and years. The credit_fund variable has the largest share of missing values (39%). Zero values of this indicator in the data source mean that no data are available, not that the lending volume was zero. Such values are treated as missing (NA) rather than as zero lending. No imputation is performed in the econometric models: each model is estimated on the observations for which all of its variables are available. The variable credit_fund therefore enters only the component-level robustness checks and the Random Forest, while the main models use the aggregate credit_total. The only exception to the no-imputation rule is the cluster analysis (Section 4.4), which requires a complete set of regional characteristics; its sensitivity to this choice is examined in Section 5.3.

The land_share variable has 14.9% missing values, owing to the specifics of registering land rights. To retain as many observations as possible, the regression analysis uses the

aggregate credit_total indicator, which is available for 170 of the 175 observations (2.86% missing). Missing values are mainly of institutional and administrative origin (Appendix A, Table A1).

The variables, their definitions, units of measurement and data sources are described in Table 1.

Table 1. Description of variables.

Variable	Description	Unit of Measurement	Source	Transformation
Dependent variable				
women_farms	Number of registered peasant farms headed by women	Number	Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan. www.stat.gov.kz	ln(1 + x)
Independent variables				
credit_total	Total amount of loans issued to women entrepreneurs in agriculture through "KazAgroFinance" JSC, "Agrarian Credit Corporation" JSC and "Fund for Financial Support of Agriculture" JSC.	Million tenge	Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan. www.stat.gov.kz	ln(1 + x)
credit_kaf	Volume of loans issued through "KazAgroFinance" JSC	Million tenge	Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan. www.stat.gov.kz	ln(1 + x)
credit_acc	Volume of loans issued through "Agrarian Credit Corporation" JSC	Million tenge	Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan. www.stat.gov.kz	ln(1 + x)
credit_fund	Volume of loans issued through "Fund for Financial Support of Agriculture" JSC	Million tenge	Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan. www.stat.gov.kz	ln(1 + x)
Independent variable (institutional)				
land_share	Percentage of women granted agricultural land: national SDG indicator 5.a.1, "Proportion of women granted agricultural land, by form of ownership" (total population), a modified version of the global indicator. Its numerator, denominator and reference period are not specified in the national metadata (Section 4.1); the published values are used without recalculation.	%	Bureau of National Statistics, SDG indicator 5.a.1 (https://stat.gov.kz/ru/sustainable-development-goals/goal/5, accessed on 17 September 2026); administrative data of the Land Management Committee of the Ministry of Agriculture of the Republic of Kazakhstan and local executive bodies.	Without transformation
Identifiers				
region_id	Region identifier (1–20)	–	Assigned by the authors	Without transformation
year	Year of observation	Year	From source data	Without transformation

4.2. Conceptual Model and Analytical Strategy

The conceptual model of the study is based on two theoretical approaches: financial (access to credit) and institutional (access to land).

Several methods are used to test the hypotheses. Pooled OLS panel regression and the TWFE model are used to assess the relationship between agricultural lending, the published land_share indicator, and the number of women-led farms. Pooled OLS captures both between-region and within-region variation, while TWFE captures changes within regions over time. These models are used to test H1 and H2.

K-means cluster analysis is used to test H3. It divides regions into groups with similar values of the indicators studied. In addition, Random Forest is used to evaluate the predictive ability of the factors under consideration and to check for possible nonlinear relationships.

The conceptual model illustrating the relationships examined in H1–H3 is presented in Figure 1.

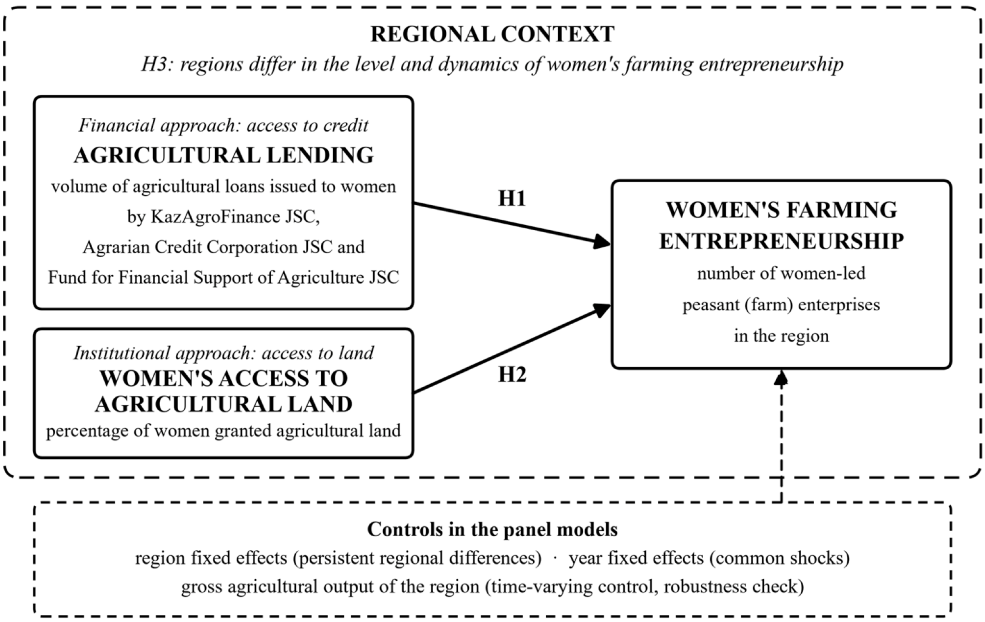


Figure 1. Conceptual model of the study: agricultural lending, women's access to agricultural land, and women's farming entrepreneurship in the context of regional heterogeneity. Note: Solid arrows represent hypothesized statistical associations (H1, H2), not causal effects; the dashed frame denotes the regional context, examined through descriptive cluster analysis (H3); the dashed box lists the controls used in the panel models.

The research scheme, which reflects the sequence of analytical steps, is shown in Figure 2.

4.3. Econometric Model: Testing H1 and H2

To test H1 and H2, Pooled OLS is first estimated as a baseline specification, and the main estimation is then performed using the TWFE model.

The Pooled OLS model estimates the overall conditional relationship between the variables by pooling observations across regions and years, without region fixed effects (Equation (1)):

$$\log_women_farms_{it} = \beta_0 + \beta_1 \cdot \log_credit_total_{it} + \beta_2 \cdot land_share_{it} + \varepsilon_{it} \quad (1)$$

Because it pools all observations, the Pooled OLS model cannot distinguish between persistent interregional differences and changes occurring within regions over time.

To separate these two sources of variation explicitly, we additionally estimate two models. The first is a between estimator, i.e., OLS on the region means of all variables, which

uses only differences between regions. The second is a two-way Mundlak regression [58,59]: pooled OLS of log_women_farms on the explanatory variables, their region means, year dummies and the region means of the year dummies, with standard errors clustered by region. In an unbalanced panel, this specification reproduces the TWFE coefficients exactly (the within-region association). The sum of the coefficients on a variable and on its region mean gives the between-region association, and a test that the coefficient on the region mean is zero tests whether the two associations differ.

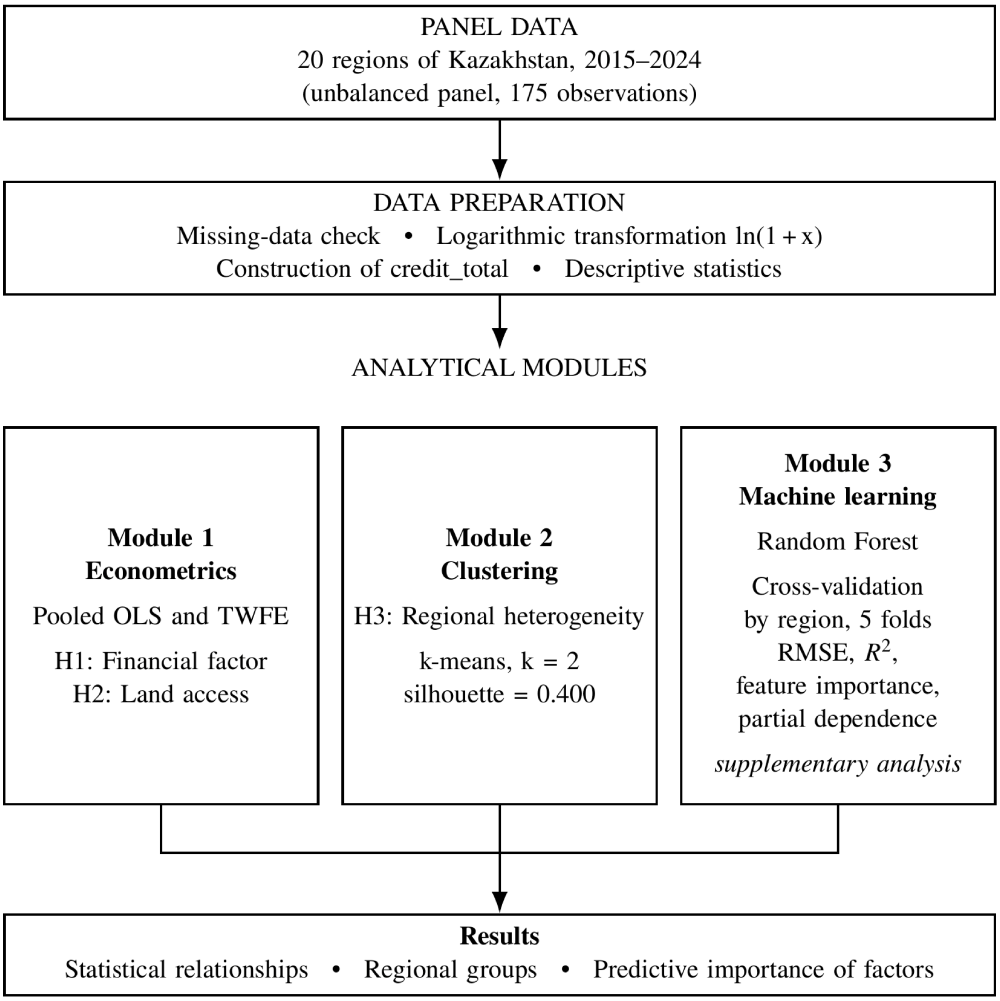


Figure 2. Research scheme: the sequence of analytical steps.

To account for regional characteristics that are constant over time (geography, infrastructure, institutional conditions) and common shocks (macroeconomic fluctuations, policy changes), we use the TWFE model (Equation (2)):

$$\log_women_farms_{it} = \alpha_i + \lambda_t + \beta_1 \cdot \log_credit_total_{it} + \beta_2 \cdot land_share_{it} + \varepsilon_{it} \quad (2)$$

where:

$\log_women_farms_{it}$—logarithm of the number of women-led farms in region i in year t.
α_i—region fixed effect.
λ_t—year fixed effect.
β_1, β_2—coefficients of the explanatory variables.
ε_{it}—error term.

The TWFE model was chosen for three reasons: (1) it accounts for unobserved regional heterogeneity; (2) it controls for common time shocks through year fixed effects; and (3) it corresponds to the structure of the regional panel sample.

TWFE controls for regional characteristics that remain constant over time and for common temporal changes. However, the recent literature emphasizes that TWFE estimates should not be automatically interpreted as evidence of causation, especially when effects are heterogeneous across regions or over time or when identification assumptions are violated [31–33]. In this study, TWFE coefficients are interpreted as conditional statistical relationships rather than causal effects.

Standard errors in the TWFE model are clustered by region. The estimation sample of the main model contains only 17 regional clusters (the three cities of national significance drop out because land_share is not available for them). Conventional cluster-robust inference can be unreliable when the number of clusters is small [60,61]. We therefore additionally report wild cluster bootstrap *p*-values for the TWFE coefficients (Rademacher weights, 9999 replications, null hypothesis imposed).

To check the robustness of the results, we replaced the aggregate lending indicator with three components: loans from "KazAgroFinance" JSC (log_credit_kaf), "Agrarian Credit Corporation" JSC (log_credit_acc), and the "Fund for Financial Support of Agriculture" JSC (log_credit_fund), and estimated a separate regression for each component. The components were not included together with log_credit_total because they are part of the aggregate indicator.

In addition, the robustness of the main TWFE result is assessed in four ways. (1) A time-varying regional control is added: the logarithm of the gross agricultural output of the region (million tenge, current prices [62]), which captures changes in the scale of regional agriculture that may affect both lending volumes and the number of women-led farms. The common price level is absorbed by the year fixed effects. (2) One- and two-year lags of total lending are added. (3) The three regions formed in 2022 (Abai, Zhetysu and Ulytau), which have only three years of data, are excluded. (4) land_share is dropped, which returns the three cities of national significance to the sample.

4.4. Cluster Analysis of Regional Heterogeneity

To test H3, we applied cluster analysis using the k-means algorithm [63].

Clustering is based on six characteristics that reflect the level and dynamics of women's farming entrepreneurship, as well as the financial and institutional context of each region:

1. Mean of log_women_farms;
2. Standard deviation of log_women_farms;
3. Mean of log_credit_total;
4. Standard deviation of log_credit_total;
5. Mean of land_share;
6. Linear time trend of log_women_farms.

The mean of land_share summarizes the published administrative indicator over the observed years. For the three cities of national significance (Astana, Almaty and Shymkent), land_share is not available; in the cluster analysis their average land_share is set to zero, which is the only case of value substitution in the study.

Because the characteristics are measured on different scales, they were standardized before clustering (z-scores: the mean was subtracted and the result was divided by the standard deviation). The optimal number of clusters was determined using the silhouette coefficient, which was highest for the two-cluster solution among the variants considered ($k = 2$; silhouette = 0.400).

The stability of the resulting cluster structure was additionally evaluated by re-running the k-means algorithm with 100 different initial centers. Its stability with respect to the composition of the sample was assessed by leaving out one region at a time, by re-clustering 1000 random subsamples of 16 of the 20 regions, and by re-clustering without the three cities. Agreement with the original partition is measured by the adjusted Rand index (ARI) [64], and the stability of each cluster by the mean cluster-wise Jaccard similarity over the subsamples [65]. Cluster analysis was used only to describe regional heterogeneity, not to establish causal relationships.

4.5. Machine Learning (Random Forest) as a Supplementary Analysis

In addition to the econometric analysis, we applied the Random Forest algorithm [66,67] for two purposes: (1) to estimate the relative predictive importance of the credit variables and land_share; and (2) to identify possible nonlinear relationships that linear regression models may not detect.

The Random Forest model was trained to predict log_women_farms based on the following features: log_credit_total, log_credit_kaf, log_credit_acc, log_credit_fund, lag1_log_credit_total, land_share and year.

The model was estimated in scikit-learn with fixed, conventional hyperparameters rather than a tuned configuration: 100 trees, a maximum tree depth of 10, at least 5 observations required to split a node and at least 2 observations per leaf, with all features considered at each split. Because the Random Forest serves an exploratory purpose and its out-of-sample performance is weak (see Section 5.4), the hyperparameters were not tuned further.

To evaluate the model's ability to generalize while preserving the panel structure of the data, we used group cross-validation (GroupKFold) with five folds defined by the region identifier.

This approach ensures that observations from the same region never appear in both the training and the test sets, which gives a more realistic assessment of the model's ability to generalize to regions not represented in the training set. Prediction accuracy was evaluated using the root-mean-square error (RMSE) and the coefficient of determination (R^2) on the test folds.

We used two approaches to interpret the model. The first is impurity-based feature importance, i.e., the reduction in mean squared error achieved by splits on each feature, averaged over all trees. The second uses partial dependence plots for the key variables (lending and land_share), which show how the predicted value of log_women_farms changes with a single variable when averaged over the other features.

The Random Forest results are supplementary and do not replace the econometric analysis. The method is used to assess the predictive ability of the factors and to identify possible nonlinear relationships; its results are not used to establish causal relationships.

4.6. Software

All calculations were performed in Python 3.14; the analysis was first carried out in Python 3.13 and then re-run in full, with virtually identical results. Pooled OLS and TWFE panel models were estimated with statsmodels 0.15 and linearmodels 7.0 (PanelOLS). K-means clustering, the Random Forest model, partial dependence, group cross-validation and feature standardization were implemented with scikit-learn 1.9, and wild cluster bootstrap *p*-values were computed with the wildboottest package (version 0.3). Data handling relied on pandas, numpy and openpyxl, with scipy for auxiliary statistics, and figures were produced with matplotlib and seaborn.

5. Results

This section presents the results of the empirical analysis of the relationship between financial and institutional factors and women's farming entrepreneurship in the regions of Kazakhstan. It first describes women-led farms and the related lending and land_share indicators across regions. It then reports the panel regression results for H1 and H2 and the cluster analysis results for H3 and concludes with the Random Forest analysis of the predictive importance of the factors.

5.1. Characteristics of Women-Led Farms in the Regions of Kazakhstan

Descriptive statistics of the key variables are presented in Table 2.

Table 2. Descriptive statistics of key variables.

Variable	N	Mean	Std. Dev.	Min	Median	Max
women_farms	175	3127.45	4298.91	14.00	1440.00	20,551.00
credit_total (million tenge)	170	3983.52	6236.45	7.80	2641.15	73,059.00
credit_kaf (million tenge)	155	905.43	1134.92	4.00	428.00	5597.00
credit_acc (million tenge)	153	2726.15	6281.00	0.00	1201.00	72,963.00
credit_fund (million tenge)	107	1119.20	630.52	129.30	941.00	2811.00
land_share (%)	149	2.01	2.14	0.01	1.19	12.10
log_women_farms	175	7.34	1.24	2.71	7.27	9.93
log_credit_total	170	7.76	1.14	2.17	7.88	11.20

Note: Logarithmic variables are calculated as $\ln(1 + x)$ and are used in the econometric analysis. The original values are shown to make the economic scale of the indicators easier to interpret.

The average number of women-led farms is 3127, ranging from 14 to 20,551, which indicates large differences between regions. Total lending to women entrepreneurs ranges from 7.8 to 73,059 million tenge, with a mean of 3983.5 million tenge. The land_share indicator averages 2.01% (median 1.19%) and ranges from 0.01% to 12.10% across regions and years.

Figure 3 shows the dynamics of the average number of women-led farms in the regions of Kazakhstan. The average increased from about 2400 farms in 2015 to about 3700 in 2024, which indicates growth in women's farming entrepreneurship.

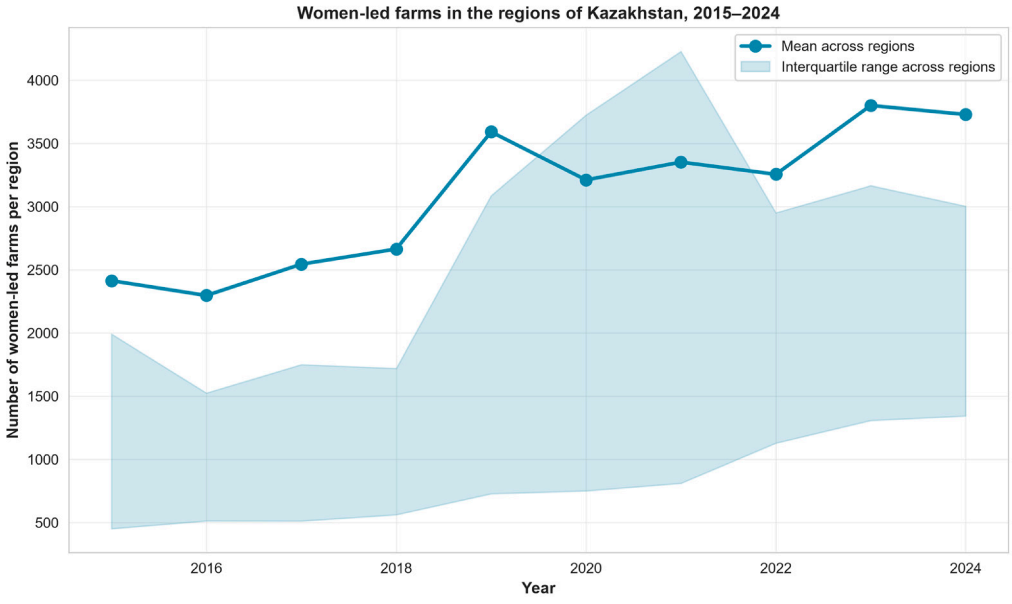


Figure 3. Dynamics of the average number of women-led farms in the regions of Kazakhstan, 2015–2024. The line shows the mean across regions; the shaded area shows the interquartile range.

The average volume of loans issued to women entrepreneurs in agriculture also increased, from about 1.4 billion tenge per region in 2015 to 4.4 billion tenge in 2023 and 9.4 billion tenge in 2024 (Figure 4). The 2024 value is raised by Astana (73.1 billion tenge); without it, the average is 6.1 billion tenge.

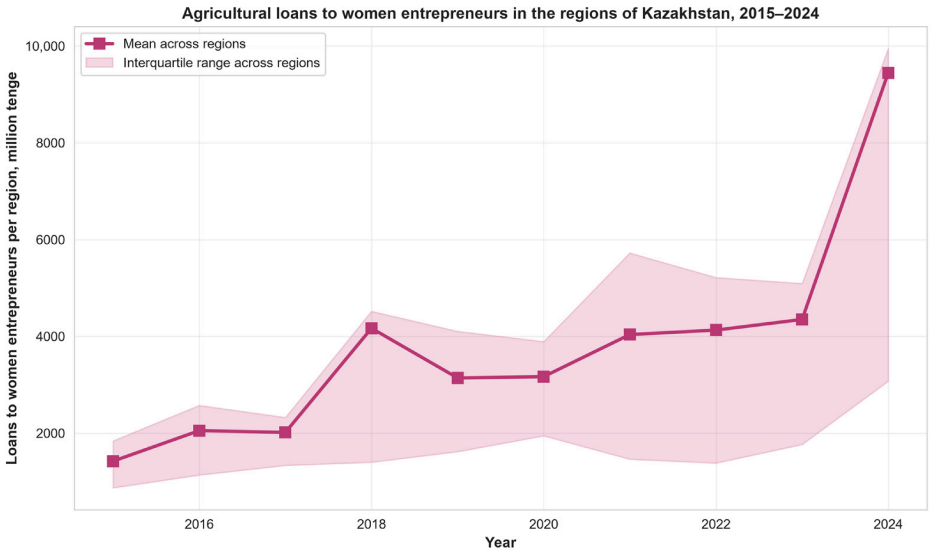


Figure 4. Dynamics of the average volume of loans issued to women entrepreneurs in agriculture per region of Kazakhstan, 2015–2024. The line shows the mean across regions; the shaded area shows the interquartile range.

The correlation analysis in Figure 5 shows a positive relationship between the volume of lending and the number of women-led farms ($r = 0.576$, $p < 0.001$). Regions with more lending tend to have more such farms. However, the correlation does not account for time-invariant regional characteristics or common time factors, so it does not indicate a stable intraregional relationship. The Pooled OLS and TWFE results are presented in Section 5.2.

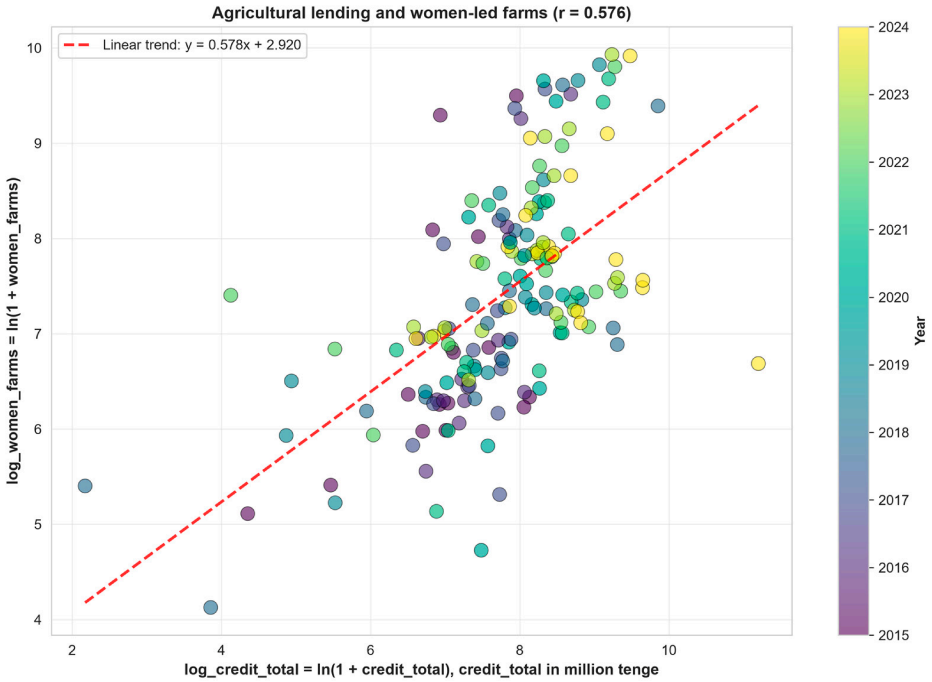


Figure 5. Relationship between the loan volume (log_credit_total) and the number of women-led farms (log_women_farms).

5.2. Panel Regression Results: Testing H1 and H2

To test H1 and H2, we move from correlation analysis to panel regression. Table 3 compares the Pooled OLS and TWFE estimates. After excluding observations with missing lending or land_share data, the estimation sample of the main TWFE model contains 149 observations.

Table 3. Comparison of Pooled OLS and TWFE estimates.

Variable	Pooled OLS		TWFE	
	Coefficient	**Std. Error**	**Coefficient**	**Std. Error**
log_credit_total	0.7250 ***	(0.0901)	0.0443	(0.0834)
land_share	0.0566 *	(0.0341)	0.0181	(0.0289)
Constant	1.7143 **	(0.6916)	—	—
Wild cluster bootstrap *p*-values (TWFE)				
log_credit_total	—		0.606	
land_share	—		0.529	
Model Statistics				
N	149		149	
Regions (clusters)	17		17	
R^2	0.336		0.030 (within)	
Adjusted R^2	0.327		—	
Overall R^2	—		0.099	
Fixed effects	No		Region + Year	
SE type	HC1-robust		Clustered (region)	

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Wild cluster bootstrap: Rademacher weights, 9999 replications, null hypothesis imposed, clustering by region.

5.2.1. Pooled OLS and Between-Region Results

In the Pooled OLS model, the coefficient on lending (log_credit_total) is positive and statistically significant (0.725; $p < 0.001$). The model explains 33.6% of the variation in the dependent variable in the pooled sample. The coefficient on land_share is also significant at the 10% level (0.0566; $p = 0.097$), but it is not statistically significant in the TWFE model (0.0181; $p = 0.531$). Because Pooled OLS combines differences between regions with changes within regions over time, its coefficient reflects an overall association rather than a purely interregional one; the two components are separated in Table 4.

Table 4. Between-region and within-region associations.

	log_credit_total				**land_share**			
Estimate	**Coef. (SE)**	**95% CI**	***p***	**Wild *p***	**Coef. (SE)**	**95% CI**	***p***	**Wild *p***
Pooled OLS (between + within)	0.725 (0.090)	[0.548, 0.902]	<0.001	—	0.057 (0.034)	[−0.010, 0.123]	0.097	—
Between estimator (region means)	0.798 (0.339)	[0.134, 1.462]	0.018	—	0.093 (0.110)	[−0.122, 0.309]	0.396	—
Between effect (two-way Mundlak)	0.953 (0.370)	[0.228, 1.677]	0.010	0.070	0.066 (0.096)	[−0.123, 0.255]	0.491	0.479
Within effect (TWFE)	0.044 (0.083)	[−0.121, 0.209]	0.596	0.606	0.018 (0.029)	[−0.039, 0.075]	0.531	0.529
Between minus within (Mundlak test)	0.909 (0.355)	[0.213, 1.604]	0.010	0.046	0.048 (0.102)	[−0.151, 0.248]	0.635	0.644

Note: N = 149 observations, 17 regions; the between estimator uses the 17 region means (HC1-robust SE). Two-way Mundlak regression: pooled OLS with region means of the regressors, year dummies and region means of the year dummies; SE clustered by region; its within effect coincides with the TWFE estimate. The 95% CI: confidence interval based on the conventional standard errors of each model (normal approximation; t distribution with the residual degrees of freedom for TWFE), not obtained by inverting the wild cluster bootstrap test. Wild *p*: wild cluster bootstrap (Rademacher weights, 9999 replications, null hypothesis imposed). —: not computed; the wild cluster bootstrap is reported only for the models with standard errors clustered by region.

The between estimator, which uses only the 17 region means, gives a positive coefficient for lending of 0.798 (95% CI 0.134 to 1.462; $p = 0.018$), and the between-region effect from

the two-way Mundlak regression is 0.953 (95% CI 0.228 to 1.677; $p = 0.010$). However, the wild cluster bootstrap p-value of the Mundlak between-region effect is 0.070, which does not meet the conventional 5% significance threshold. The within-region effect equals the TWFE estimate of 0.044 (95% CI −0.121 to 0.209; $p = 0.596$). The difference between the two is 0.909 and is statistically significant ($p = 0.010$; wild cluster bootstrap $p = 0.046$). The positive pooled association between lending and the number of women-led farms therefore appears to reflect mainly differences between regions, while no statistically significant association is detected within regions over time. Because the between estimator uses only 17 regional means and the Mundlak between-region effect is not significant at the 5% level under the wild cluster bootstrap, the between-region evidence is imprecise and should be interpreted with caution. For land_share, neither the between-region nor the within-region association is statistically significant. Thus, H1 receives only limited empirical support at the interregional level.

5.2.2. TWFE Results: Intraregional Level

In the TWFE model, the coefficient on lending decreases from 0.725 to 0.044 and is no longer statistically significant ($p = 0.596$; 95% CI −0.121 to 0.209). The coefficient on land_share is not significant either ($p = 0.531$; 95% CI −0.039 to 0.075). Thus, no statistically significant association between lending and the number of women-led farms is detected within regions over time. This does not show that no within-region relationship exists: the confidence interval includes both negative and positive values. However, its upper limit (0.209) is well below the lower limit of the Pooled OLS interval (0.548), so the data do not support a within-region association as large as the pooled estimate.

Because the loss of significance of the credit coefficient is central to this conclusion, we checked that it does not depend on the small number of clusters (17 regions). The wild cluster bootstrap gives $p = 0.606$ for log_credit_total and $p = 0.529$ for land_share (Table 3), very close to the conventional cluster-robust p-values (0.596 and 0.531). The absence of a statistically significant within-region relationship is therefore robust to the inference method.

Figure 6 shows how the lending coefficient changes from Pooled OLS to TWFE. In Pooled OLS, the coefficient is positive and significant (0.725, $p < 0.001$), whereas in TWFE it decreases to 0.044 and is not statistically significant. This large decrease after the inclusion of region and year fixed effects is consistent with the between-region results in Table 4.

In summary, the positive association between the volume of lending and the number of women-led farms found in the pooled and between-region estimates is not confirmed by the analysis of intraregional dynamics. On average, regions with more lending have more women-led farms, but growth in lending within a region is not accompanied by a statistically significant increase in their number. H1 therefore receives partial support: the positive association between agricultural lending and women's farming is observed mainly at the interregional level, while no statistically significant association is detected after region and year fixed effects are taken into account.

H2 does not receive empirical support from the available measure: within regions, changes in land_share are not significantly related to changes in the number of women-led farms ($p = 0.531$ in the TWFE model).

5.2.3. Robustness Checks

In the first robustness check, the aggregate lending indicator was replaced with its three components: loans from "KazAgroFinance" JSC (log_credit_kaf), "Agrarian Credit Corporation" JSC (log_credit_acc) and the "Fund for Financial Support of Agriculture" JSC (log_credit_fund). The results are presented in Table 5. None of the three sources

shows a statistically significant relationship with the dynamics of women-led farms in the TWFE model: the coefficients are small (from −0.0473 to 0.0031), and all p-values exceed 0.05 (0.933, 0.331 and 0.910). Because of missing data on the Fund for Financial Support of Agriculture, the sample is reduced to 88 observations. To ensure comparability of the coefficients, the models for the three credit sources were estimated on the same sample of 88 observations, for which data on all three components are available.

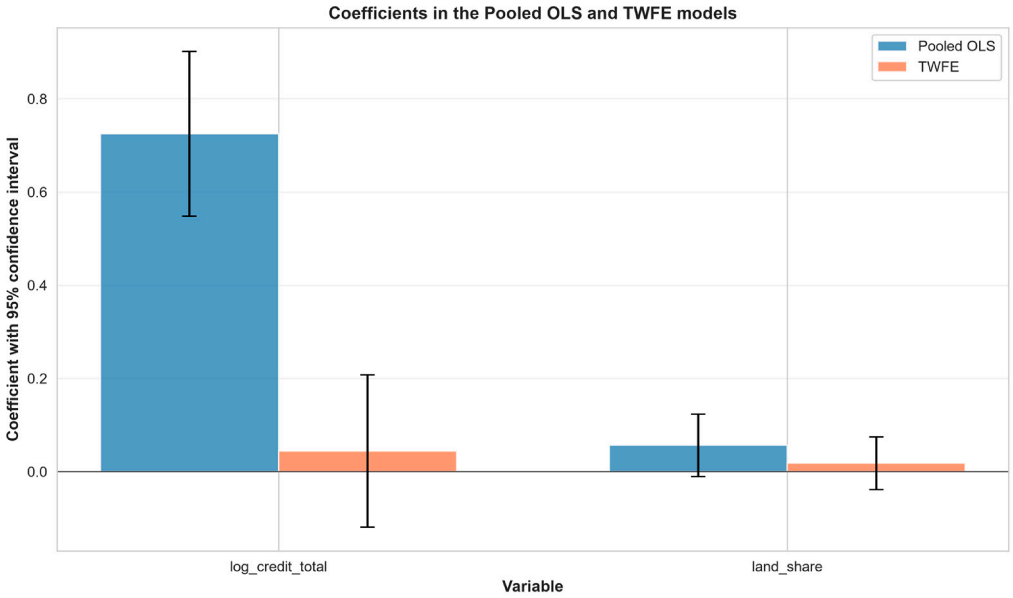


Figure 6. Comparison of the Pooled OLS and TWFE estimates. Error bars show 95% confidence intervals (HC1-robust standard errors for Pooled OLS, standard errors clustered by region for TWFE).

Table 5. TWFE estimates by lending source.

Variable	Coefficient	Standard Error	t-Statistic	p-Value	N
log_credit_kaf	0.0031	(0.0365)	0.084	0.933	88
log_credit_acc	−0.0473	(0.0484)	−0.979	0.331	88
log_credit_fund	−0.0073	(0.0647)	−0.113	0.910	88

Table 6 reports the further robustness checks described in Section 4.3. Adding the gross agricultural output of the region as a time-varying control does not change the conclusion: the credit coefficient becomes −0.041 ($p = 0.460$; wild cluster bootstrap $p = 0.486$), while gross agricultural output itself is positively associated with the number of women-led farms within regions (1.038; $p < 0.001$). Neither current nor lagged lending is statistically significant when one- and two-year lags are added. Excluding the three regions formed in 2022 leaves the coefficient practically unchanged (0.047; $p = 0.585$). However, excluding the newly established regions does not fully address the comparability issue, as the boundaries of their parent regions (Almaty, East Kazakhstan, and Karaganda) also changed in 2022. Controlling for gross agricultural output may account for some changes in regional agricultural scale but cannot fully eliminate the resulting structural break.

The credit coefficient is significant under conventional cluster-robust standard errors only in the two specifications that drop land_share and thereby return the three cities of national significance (Astana, Almaty and Shymkent) to the sample (0.118, $p = 0.020$; with the gross output control, 0.110, $p = 0.027$). These results depend on the sample and the inference method: with the same 17 regions but without land_share, the coefficient is 0.042 ($p = 0.606$), so the difference is driven by the three cities, where agriculture is

marginal, and neither city-inclusive estimate is significant at the 5% level under the wild cluster bootstrap ($p = 0.062$ and 0.122). Overall, no statistically significant within-region association between lending and the number of women-led farms is detected in any of the specifications estimated on the 17 regions with land_share data.

Table 6. Robustness of the within-region association of lending (TWFE).

Specification	Coef. (SE)	*p*	Wild *p*	N	Regions
Baseline (Table 3)	0.044 (0.083)	0.596	0.606	149	17
+log gross agricultural output	−0.041 (0.055)	0.460	0.486	149	17
+lag 1 of credit: current/lag 1	0.072/0.010	0.481/0.867	—	132	17
+lags 1 and 2: current/lag 2	0.031/−0.117	0.781/0.295	—	115	17
Excluding regions formed in 2022	0.047 (0.086)	0.585	—	140	14
Without land_share, cities excluded	0.042 (0.082)	0.606	—	149	17
Without land_share, cities included	0.118 (0.050)	0.020	0.062	170	20
Without land_share, cities included + log gross agricultural output	0.110 (0.049)	0.027	0.122	170	20

Note: dependent variable log_women_farms; all models include region and year fixed effects, SE clustered by region. Coefficients on log_credit_total (for the lag models, the current value and the longest lag are shown). Wild *p*: wild cluster bootstrap (Rademacher weights, 9999 replications, null hypothesis imposed); —not computed.

A relative indicator, the share of women-led farms in the total number of active peasant (farm) enterprises, could not be used as an alternative dependent variable for the full period: the total number of active farms by region is published only from 2022 onwards [68]. For 2022–2024, women-led farms account for 24–40% of all active farms (on average 29%), and the logarithms of the two counts are almost perfectly correlated (r = 0.991). The number of women-led farms therefore largely reflects the scale of the regional farm sector. In the TWFE model, persistent differences in scale are absorbed by the region fixed effects, and changes in scale are partly controlled for by gross agricultural output (Table 6).

5.3. Cluster Analysis Results: Testing H3

To test H3, k-means clustering was used to identify groups of regions that are similar in terms of women's farming entrepreneurship, lending, and the published land_share indicator, after which the clusters were compared by the rate of change in the number of women-led farms. The optimal number of clusters (k = 2) was determined by the maximum silhouette coefficient (0.400), which indicates moderate rather than strong cluster separation. Cluster profiles are shown in Table 7, and the full list of regions by cluster is given in Appendix B.

The two-cluster solution is stable with respect to the choice of initial centers: in 97 of 100 re-runs with different random seeds, the same partition is obtained (mean ARI 0.99). It is moderately stable with respect to the composition of the sample: when one region is left out, the mean ARI is 0.87, and over 1000 subsamples of 16 regions, it is 0.76. The mean cluster-wise Jaccard similarity is 0.93 for Cluster 0 and 0.78 for Cluster 1, above the 0.75 level usually regarded as indicating a stable cluster [65]. However, the partition depends substantially on the three cities, for which land_share is not available and is set to zero: without them, two clusters still give the highest silhouette coefficient (0.288), but the grouping of the remaining 17 regions is different (7 of the 17 regions change cluster). The clusters should therefore be read as a description of regional heterogeneity rather than as a robust typology of regions.

Cluster analysis revealed two groups of regions.

Cluster 0 (15 regions) has higher average levels of women-led farms (average log_women_farms 7.71), lending (8.02) and land_share (2.30%). On the original scale, this corresponds to approximately 2230 farms and 3030 million tenge of lending per region.

The average time trend coefficient is 0.091, which corresponds to an annual growth rate of about 9.5%.

Cluster 1 (5 regions) has lower average levels of women-led farms (average log_women_farms 6.35, about 570 farms per region), lending (6.75, about 860 million tenge) and land_share (0.18%). Its time trend coefficient is higher (0.296), which corresponds to an annual growth rate of about 34.5%.

Table 7. Regional cluster profiles.

Cluster	N Regions	log_women_farms (Average Value)	log_credit_total (Average Value)	Percentage of Women Granted Agricultural Land (Average Value, %)	Growth Trend (Annual)
Cluster 0	15	7.71	8.02	2.30	0.091
Cluster 1	5	6.35	6.75	0.18	0.296

Note: The table shows the average values of the indicators for each cluster. "Growth trend" is the coefficient of a linear time trend in log_women_farms, which reflects the annual increase in the number of women-led farms. Cluster means of land_share include the zero values assigned to the three cities of national significance, for which the indicator is not available, and are not directly comparable measures of access to land.

In the scatter plot (Figure 7), Cluster 0 (purple) lies in the area of high values on both the lending and the women-led farms axes, and Cluster 1 (yellow) lies in the area of low values on both axes.

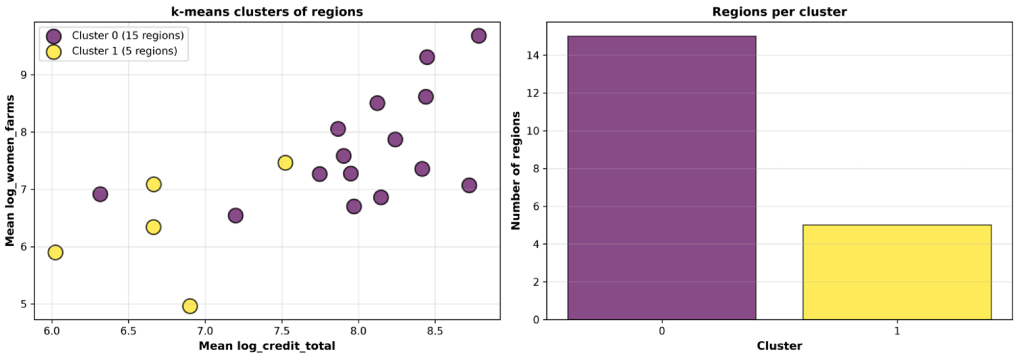


Figure 7. Distribution of regions by lending level and number of women-led farms, by cluster.

In summary, the results of the cluster analysis provide descriptive evidence of regional heterogeneity of women's farming entrepreneurship, in line with H3, rather than an unequivocal confirmation of the hypothesis. The separation between the clusters is moderate (silhouette coefficient 0.400), and the partition is sensitive to the inclusion of the three cities. The clusters differ not only in the levels of women-led farms, lending and land_share but also in the growth rate of the number of women-led farms.

Cluster 0 includes large agricultural regions such as Akmola, Kostanay, and Turkestan regions. Cluster 1 includes cities of national significance (Astana, Almaty, Shymkent) and regions with a smaller scale of agricultural production (Mangystau and Kyzylorda regions). For a complete list of regions by cluster, see Appendix B.

5.4. Random Forest Results (Supplementary Analysis)

In addition to the econometric analysis, we used Random Forest to assess the predictive importance of the factors. The model was evaluated with group cross-validation at the regional level (GroupKFold, 5 folds), which ensures that data from the same region do not appear in both the training and the test sets. The feature importances are presented in Table 8.

Table 8 shows that loans from the Fund for Financial Support of Agriculture (log_credit_fund, 44.8%) make the largest contribution to the predictive ability of the

model. However, this result should be interpreted with caution. The variable has a large share of missing values (39%) and is structurally related to the aggregate indicator of total lending. The high importance of this feature does not indicate that this source of funding has a stronger impact on the development of women-led farms. Including credit_fund also restricts the Random Forest sample to 75 observations. Re-estimating the model without credit_fund (122 observations) does not improve its out-of-sample performance, which suggests that the high importance of credit_fund partly reflects the pattern of missing data rather than a stable predictive signal.

Table 8. Feature importance in the Random Forest model.

Feature	Importance	% of Total
log_credit_fund	0.448	44.77%
log_credit_total	0.146	14.56%
log_credit_acc	0.102	10.23%
lag1_log_credit_total	0.102	10.22%
land_share	0.088	8.84%
log_credit_kaf	0.082	8.19%
year	0.032	3.19%

The variable log_credit_total has a relative importance of 14.6%, while log_credit_acc and the one-year lag of total lending (lag1_log_credit_total) each have about 10.2%. The relative importance of land_share is 8.8%, and that of log_credit_kaf is 8.2%. Taken together, the financial indicators account for a large part of the feature importance in the Random Forest model. However, their relative importance cannot be interpreted as an estimate of a statistical or causal effect.

At the same time, the predictive ability of the model for regions not used in training is low. The negative cross-validated coefficient of determination (-1.12, mean over the five folds) indicates low predictive ability on the test data. Therefore, the Random Forest results cannot be considered independent confirmation of the econometric conclusions and are used only as a supplementary, exploratory predictive analysis.

The lag1_log_credit_total variable was also included in the model to capture a possible delayed relationship with lending. Its relative importance is 10.2%. However, the lagged variable was not included in the main TWFE model because it reduces the sample by one year, which matters in a panel with missing data. The TWFE specifications with lagged lending do not show statistically significant coefficients (Table 6).

The partial dependence plot in Figure 8 shows that the predicted log_women_farms generally increases with the volume of lending (log_credit_total).

Figure 9 shows the partial dependence of the predicted number of women-led farms (log_women_farms) on the published administrative indicator land_share. The shape of the dependence changes at approximately 2%, which may indicate a nonlinear relationship. However, this result is not statistical evidence of a threshold effect and requires verification with a formal econometric specification. A preliminary check of this kind does not support a threshold at about 2%. In TWFE models with a dummy for land_share above a threshold, estimated over a grid of thresholds between the 10th and 90th percentiles of land_share, none of the threshold dummies is statistically significant (for thresholds between 1.8% and 2.4%, p ranges from 0.48 to 0.68). The pattern in Figure 9 should therefore be treated as exploratory.

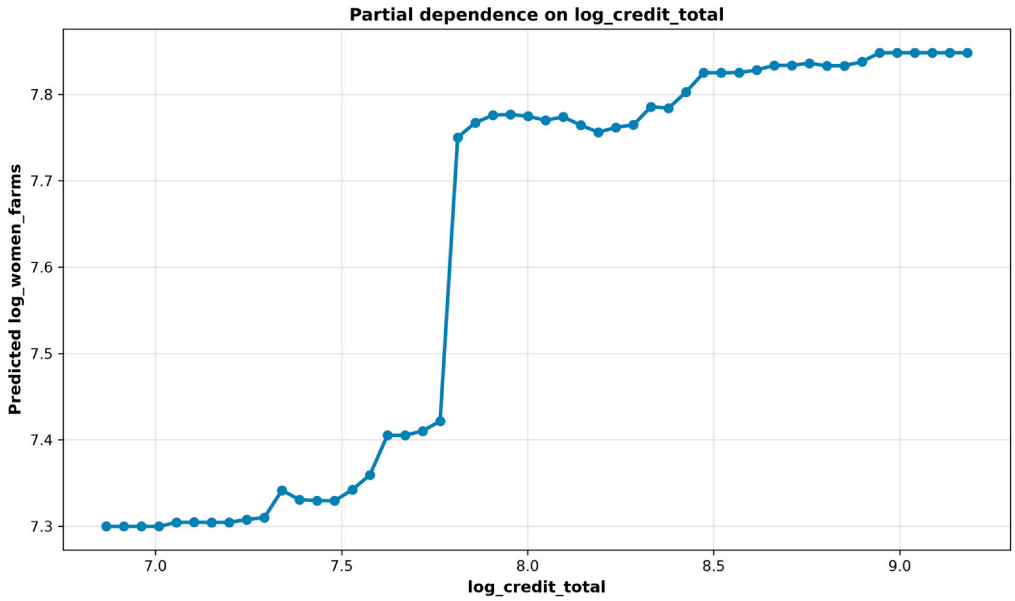


Figure 8. Partial dependence of the predicted log_women_farms on log_credit_total (Random Forest), between the 5th and 95th percentiles of log_credit_total.

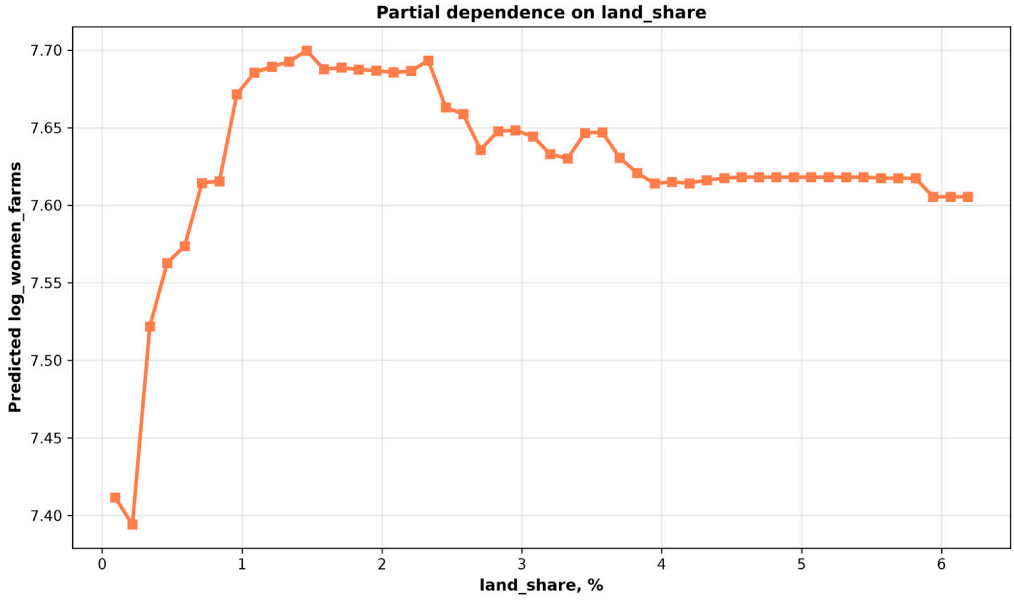


Figure 9. Partial dependence of the predicted log_women_farms on land_share (Random Forest), between the 5th and 95th percentiles of land_share.

6. Discussion

This section discusses the results in relation to the hypotheses and compares them with the findings of existing studies. The discussion is structured around three main themes: (1) financial factors and women-led farms (H1); (2) institutional factors and access to land (H2); and (3) regional heterogeneity (H3). The section then compares the results of the different methods and concludes with the theoretical and policy implications, the limitations of the study and directions for future research.

6.1. Agricultural Lending and Women-Led Farms: H1

The key result of the study is the discrepancy between the Pooled OLS and TWFE estimates. In the pooled model, the lending coefficient is positive and significant

($\beta = 0.725$; $p < 0.001$), but in the TWFE model it decreases to 0.044 and is not statistically significant ($p = 0.596$; 95% CI −0.121 to 0.209). Pooled OLS combines differences between regions with changes within regions; the between estimator and the Mundlak regression (Table 4) indicate that the pooled association is driven mainly by differences between regions, whereas the TWFE model captures changes within regions over time. The between-region evidence is itself limited: the Mundlak between-region effect is significant under conventional cluster-robust inference ($p = 0.010$) but not at the 5% level under the wild cluster bootstrap ($p = 0.070$).

The results provide limited support for H1 at the interregional level but do not show a statistically significant intraregional relationship between the volume of lending and the dynamics of women-led farms. Regions with more lending have, on average, more women-led farms, but an increase in lending within a particular region is not accompanied by a statistically significant increase in their number.

These results are consistent with studies that emphasize the importance of financial resources for the development of women's entrepreneurship [34–39,69]. However, they also indicate that this relationship depends on the level of analysis. Whereas studies in Africa and South Asia often associate microfinance with improved business performance [4–8], in Kazakhstan the relationship appears mainly at the interregional level.

One possible explanation lies in differences in institutional and regional contexts. In Kazakhstan, access to lending may depend not only on the immediate availability of financial resources but also on the existing economic infrastructure, entrepreneurial potential, and regional characteristics. This explanation is consistent with studies showing that financial constraints are combined with infrastructural, institutional, and resource constraints [2,40–44,70].

The results for individual lending sources also do not show a statistically significant intraregional relationship: none of the three sources has a significant coefficient in the TWFE model. This means that, whatever the source, changes in the volume of lending within a region are not significantly related to changes in the number of women-led farms.

Overall, the results suggest that the positive association between lending and the development of women-led farms mainly reflects differences between regions. Regions with higher levels of economic and entrepreneurial activity also have a greater volume of lending and a greater number of women-led farms. Therefore, a positive interregional relationship does not mean that increased lending leads to an increase in the number of women-led farms.

6.2. Institutional Factors and Access to Land: H2 and Measurement Limitations

In the TWFE model, there was no statistically significant intraregional relationship between land_share and the number of women-led farms ($p = 0.531$). Thus, H2 is not supported by the available measure.

Research suggests that women's access to agricultural land may support entrepreneurship through several potential mechanisms. Secure land rights may facilitate access to credit, including through the use of land as collateral, and create incentives for long-term agricultural investment [2]. In addition, differences between documented ownership, transfer rights, and actual decision-making authority may affect women's ability to control and use agricultural land [10].

Several factors may explain the absence of a statistically significant within-region relationship. First, land_share is an administrative indicator whose numerator and denominator are not documented (Section 4.1). Its values change abruptly in some regions (for example, from 12.10% to 0.29% in Kostanay region between 2015 and 2016), which may reflect changes in administrative recording, although their causes are not documented. Such discontinuities

may affect the within-region estimates, and the direction of any resulting bias is unknown. In addition, being granted a land plot does not necessarily mean that a woman has full control over this resource.

Second, the within-region variation of the indicator may be limited or heterogeneous. The standard deviation of land_share is 2.14 percentage points with a mean of 2.01%. The indicator varies between observations, but a large part of this variation may reflect differences between regions. TWFE estimates the relationship from changes within regions, and if such changes are small, the model may not detect a statistically significant association.

Third, any relationship with access to land may appear only with a delay. Lagged variables could be used to test for such a delay. However, because of the limited sample size, we did not include them in the main specification.

The Random Forest partial dependence plot (Figure 9) suggests a possible nonlinear form of the predicted relationship between land_share and the number of women-led farms, with a change in shape at about 2%. This result is predictive in nature and is not statistical confirmation of a threshold effect. Testing for a threshold effect requires specialized econometric methods and may be a subject for further research. A preliminary TWFE threshold test did not support a threshold at about 2% (Section 5.4).

Thus, the absence of a statistically significant relationship does not allow us to conclude that land constraints do not matter for women farmers or that land rights are merely an indicator of other regional factors. Without a documented numerator, denominator and reference period, the level of the indicator cannot establish the extent of women's access to agricultural land or the magnitude of gender inequality in land access.

6.3. Regional Heterogeneity: Descriptive Evidence for H3

The cluster analysis identified two groups of regions that differ in the level and dynamics of women-led farms, as well as in lending and the published land_share indicator. These results provide descriptive evidence consistent with H3. However, given only 20 regions, a moderate silhouette coefficient (0.400) and the sensitivity of cluster membership to the composition of the regional sample, in particular to the three cities (Section 5.3), these groups should be interpreted as exploratory rather than stable classifications.

Cluster 0 (15 regions) has higher average levels of women-led farms, lending and land_share. Cluster 1 (5 regions) has lower average levels but a higher growth rate.

The higher growth rate in Cluster 1 despite lower average levels of women-led farms and lending suggests that differences between regions are not limited to current levels of lending; the cluster profiles do not, however, establish differences in actual land access. This result is consistent with the literature that stresses the importance of infrastructure, the institutional environment, and market access for entrepreneurship development. In some countries in Africa and South Asia, microcredit can serve as a tool for women's economic empowerment when access to traditional finance is limited. In Kazakhstan, by contrast, the effectiveness of credit support may also depend on other conditions, including the availability of collateral, infrastructure, market access and the institutional environment.

Cluster 0 includes mainly regions with developed agriculture (e.g., Akmola, Kostanay, Turkestan and Almaty regions).

Cluster 1 includes cities of national significance (Astana, Almaty, Shymkent) and regions with a smaller scale of agricultural production (Mangystau and Kyzylorda regions).

The higher growth rate in Cluster 1 (approximately 34.5% vs. 9.5% in Cluster 0) may be related to a low-base effect (a small initial number of women-led farms): such regions may show higher growth rates even when only small-scale support programs are implemented. Testing this assumption requires data at the individual farm level.

Differences between clusters should also be taken into account when developing measures to support women's farming entrepreneurship. The results do not show that uniform credit programs are ineffective, but they do indicate that the conditions and outcomes of support may vary by region. Therefore, such measures should take into account regional differences in infrastructure, market access, advisory services and other resources. Their effectiveness requires a separate empirical test.

6.4. Comparison of the Results of Different Methods

A comparison of the results of the different methods shows that the relationship between lending and the number of women-led farms depends strongly on the level of analysis. Pooled OLS, which combines between-region and within-region variation, shows a positive association between lending and the number of women-led farms; the between estimator and the Mundlak regression attribute it mainly to differences between regions, with inference on the between-region association sensitive to the method used (Table 4). In the TWFE model, which analyzes changes within regions over time and accounts for region and year fixed effects, no statistically significant association is detected. Regional differences should therefore be taken into account when interpreting the relationship between lending and women's farming.

The cluster analysis complements the econometric results by describing regional differences in women-led farms, lending, and the published land_share indicator. However, clustering is descriptive in nature and cannot establish causal mechanisms.

Random Forest was used as a supplementary predictive tool. Despite the relative importance of individual credit-related variables, the negative cross-validated R^2 indicates low predictive ability of the model on the test data. Therefore, the Random Forest results are not used to confirm the econometric relationships.

6.5. Theoretical Implications

The findings show the importance of considering financial and institutional approaches together when examining the relationship between agricultural lending and women's farming entrepreneurship.

Empirical contribution. The study expands the empirical evidence on the relationship between agricultural lending and women's entrepreneurship in the context of Kazakhstan and Central Asia.

Analytical contribution. The combination of Pooled OLS and TWFE with the between estimator and the Mundlak regression made it possible to separate interregional relationships from intraregional dynamics, while cluster analysis and Random Forest were used as supplementary descriptive and predictive tools.

Substantive contribution. The estimates suggest a positive association between lending and the development of women-led farms across regions, but no statistically significant association is detected within regions after accounting for region and year fixed effects. Analyses of financial support for women's farming businesses should therefore take into account the institutional and regional context.

The findings also have implications for sustainable agricultural and rural development. Previous research points to the role of gender-related barriers and access to productive resources in the adoption of sustainable agricultural practices and the development of rural livelihoods [13,16,70]. Women's economic empowerment is also relevant to household food security, although participation in farming alone does not necessarily lead to improved empowerment or food security outcomes [14,15]. In addition, women's land rights and access to livelihood resources may contribute to household well-being, poverty reduction, and more inclusive rural development [2,70]. In the context of our findings, these potential

pathways indicate that regional and institutional conditions should be considered alongside agricultural lending. However, farm resilience, food security, income sustainability, and long-term rural development outcomes were not directly estimated in this study.

6.6. Practical and Policy Implications

6.6.1. Policy Implications Based on the Empirical Findings

The empirical results indicate that the relationship between agricultural lending and women's farming entrepreneurship differs between interregional and intraregional analyses. Therefore, regional differences should be considered when evaluating agricultural support programs, and lending volumes alone should not be interpreted as evidence of program effectiveness.

Differentiated recommendations by region type

For Cluster 0 (15 regions), characterized by relatively higher levels of women-led farming and agricultural lending, policy options include financial literacy programs, advisory services, support for cooperation, improved market access, and income diversification.

For Cluster 1 (5 regions), characterized by lower average levels of women-led farming and lending but higher growth rates, policy options include improved access to financial services, targeted advisory support, and investments in infrastructure and market connectivity.

These recommendations are informed by the observed regional differences and the broader literature [40,44,49,70]. However, the effectiveness of these measures was not directly evaluated in this study. Therefore, they should be regarded as potential policy options rather than empirically validated interventions. The distinction between the two region types is itself indicative, given the sensitivity of cluster membership to the composition of the regional sample (Section 5.3).

6.6.2. General Policy Recommendations

The following recommendations reflect broader policy considerations and are not directly tested as interventions in the empirical analysis.

At the national level

Improve women farmers' access to seasonal agricultural credit programs and support the effective use of these loans through advisory services, access to markets, infrastructure and financial services.

Expand advisory and information support for women farmers in obtaining and using loans.

Improve mechanisms for women farmers' access to land resources and for the registration of land rights.

Strengthen coordination between land policy, infrastructure development and financial instruments for rural entrepreneurship.

Introduce regular monitoring of the participation of women farmers in state agricultural support programs.

The proposed policy directions are aligned with SDG 5 (gender equality) and are linked to SDGs 1, 2, 8 and 10.

6.7. Limitations of the Study

The study has a number of limitations that should be considered when interpreting the results.

Aggregated data. The use of regional data, rather than farm-level data, may have led to underestimation of local and individual factors affecting women's entrepreneurship. In addition, we analyze only the number of women-led farms (an extensive indicator), not

their productivity or sustainability. Future studies could use farm-level data and include intensive indicators (output, income, land area).

Data gaps. About 39% of observations for the credit_fund variable are missing: the indicator is not available for any region in 2022–2024 (60 of the 68 missing observations), and the remaining gaps concern the three cities of national significance in 2015–2018. The models that use credit_fund (the component-level robustness checks and the Random Forest) therefore cover only 2015–2021. These gaps follow a systematic pattern by year and region, but we cannot completely rule out biased estimates due to systematic differences in reporting.

Measurement of land access. We used the percentage of women who were granted agricultural land, which does not distinguish between ownership rights, leases, and actual use. The indicator is compiled from administrative data, and its official metadata do not specify the numerator, the denominator or the reference period; the results for H2 therefore refer to this administrative indicator and may be affected by changes in administrative recording. Land quality was also not taken into account. Future studies could examine different forms of land rights separately and take into account the characteristics of the land used.

Endogeneity of lending. The TWFE model does not completely eliminate endogeneity. Loans may be directed to regions that already have high entrepreneurial potential. In addition, general economic changes, such as changes in commodity prices, may simultaneously affect lending and the number of women-led farms. Therefore, the estimates reflect statistical relationships between the indicators and do not allow us to draw causal conclusions.

Small number of regions. The study covers only 20 regions, which limits the statistical reliability of some supplementary analyses, including the cluster analysis. Therefore, the clustering results should be used primarily to describe differences between regions. In addition, the cluster partition depends on the three cities, for which land_share is not available and is set to zero in the cluster analysis.

Administrative boundary changes. The establishment of three new regions in 2022 also altered the boundaries of their parent regions (Almaty, East Kazakhstan, and Karaganda). As a result, some regional indicators may not be fully comparable before and after 2022. Excluding the newly established regions and controlling for gross agricultural output do not fully resolve this issue. Therefore, the within-region estimates should be interpreted with caution.

Potential sources of bias include omitted time-varying regional factors, incomplete reporting of lending data, and differences in the measurement of women's access to agricultural land. The TWFE model accounts for time-invariant regional characteristics and common year effects, but does not eliminate these potential sources of bias.

6.8. Future Research Directions

The results suggest several directions for future research.

Data at the level of individual women-led farms would make it possible to take into account microeconomic factors affecting women's entrepreneurship.

Instrumental variables could be used to address endogeneity and obtain more reliable estimates of causal relationships.

An analysis of the quality and sustainability of women-led farms, rather than only their number, would make it possible to assess both the extensive and the intensive effects of financial support.

Studying the interaction between land access and credit financing could reveal synergistic effects that were not found when these factors were considered separately.

A comparative analysis with other Central Asian countries would make it possible to assess the external validity of the results and to identify common patterns and country-specific features.

7. Conclusions

This paper analyzes the relationship between institutional agricultural lending, the published administrative land indicator and the development of women-led farms in 20 regions of Kazakhstan in 2015–2024. The combination of panel econometrics, cluster analysis, and machine learning made it possible to distinguish between interregional relationships and intraregional dynamics and to supplement the econometric results with descriptive and predictive analysis.

1. Lending is positively associated with the development of women-led farms across regions, but no statistically significant association is detected within regions.

After controlling for region and year fixed effects, the within-region estimate is small and imprecise (0.044; 95% CI −0.121 to 0.209): both negative and positive associations are compatible with the data, but an association as large as the pooled estimate is not supported.

Therefore, the interregional relationship does not allow us to draw conclusions about a causal impact of lending on the development of women-led farms.

Hypothesis H1 receives only partial support. The estimates suggest a stronger association between regions than within regions: the Mundlak test rejects equality of the between-region and within-region coefficients at the 5% level (wild cluster bootstrap $p = 0.046$), although the evidence for a positive between-region association is sensitive to the inference method (wild cluster bootstrap $p = 0.070$). H1 is not supported by the within-region estimates.

2. No statistically significant intraregional relationship is detected between the land_share indicator and the dynamics of women-led farms.

This provides no supporting evidence for H2 with the available measure. However, the national definition of the indicator is not documented in detail (Section 4.1), which prevents a conclusive assessment of the broader hypothesis on women's access to agricultural land.

3. The regions of Kazakhstan are characterized by pronounced heterogeneity in the level and dynamics of women-led farms and related resource indicators.

Cluster analysis revealed two groups of regions that differ in the level and growth rate of women-led farms. These differences are likely to be associated with factors not included in the model (infrastructure, institutional and demographic characteristics). The results of the cluster analysis are consistent with H3.

4. The predictive power of credit variables is limited.

Credit variables showed relatively high feature importance in the Random Forest model, particularly log_credit_fund. However, this variable has a substantial proportion of missing values (39%). The negative cross-validated R^2 indicates limited predictive performance. Therefore, feature importance should not be interpreted as evidence of statistical significance or causal influence.

The results of the study indicate the need to move from assessing agricultural support primarily on the basis of lending volumes towards a broader assessment of the conditions under which women-led farms access and use financial resources. Such an approach should take into account regional differences and institutional constraints, including access to agricultural land.

Author Contributions: Conceptualization, B.A., T.A., Z.B., G.M., G.K. and R.S.; methodology, B.A.; software, B.A.; validation, T.A. and R.S.; formal analysis, B.A.; investigation, B.A.; resources, B.A.; data curation, R.S.; writing—original draft preparation, B.A.; writing—review and editing, B.A., T.A.,

Z.B., G.M., G.K. and R.S.; visualization, B.A.; supervision, B.A.; project administration, B.A.; funding acquisition, B.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministry of Science and Higher Education of the Republic of Kazakhstan, project AP26100879: Design and Implementation of an Inclusive Agricultural Development Model, under the competition for grant funding for scientific and/or scientific-technical projects for 2025–2027.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original data presented in the study are openly available in the following publicly available official sources: the Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan at www.stat.gov.kz; data on women-led farms and agricultural lending from "KazAgroFinance" JSC, "Agrarian Credit Corporation" JSC, and "Fund for Financial Support of Agriculture" JSC in *Women and Men of Kazakhstan. Statistical Compendium*; data on land use from the Land Resources Management Committee of the Ministry of Agriculture of the Republic of Kazakhstan, "Proportion of women who have been granted agricultural land, broken down by form of ownership" (2015–2024; series for the total population), and the Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan, SDG indicator 5.a.1 data file and metadata at https://stat.gov.kz/ru/sustainable-development-goals/goal/5/ (accessed on 17 September 2026); and gross agricultural output and the number of active peasant (farm) enterprises by region from the Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan [62,68].

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

Abbreviations

The following abbreviations are used in this manuscript:

ARI	Adjusted Rand index
H1–H3	Hypotheses 1–3
HC1	Heteroskedasticity-consistent standard errors (HC1 type)
JEL	Journal of Economic Literature classification
JSC	Joint-stock company
NA	Not available (missing data)
OLS	Ordinary least squares
RMSE	Root-mean-square error
SDG	Sustainable Development Goal
SE	Standard error
TWFE	Two-way fixed effects

Appendix A

Table A1. Missing observations by variable and year (number of regions).

Year	women_farms	credit_kaf	credit_acc	credit_fund	land_share
2015	5	6	7	6	6
2016	4	8	7	6	6
2017	4	6	8	6	6
2018	3	5	6	6	6
2019	3	6	7	3	6
2020	3	6	3	3	6
2021	3	4	6	3	6
2022	0	2	3	20	3
2023	0	1	0	20	3

Table A1. *Cont.*

Year	women_farms	credit_kaf	credit_acc	credit_fund	land_share
2024	0	1	0	20	3
Total	25	45	47	93	51

Appendix B

Table A2. Regions by cluster.

Cluster	Regions
Cluster 0 (high level)	Akmola, Aktobe, Almaty region, Atyrau, East Kazakhstan, Zhambyl, West Kazakhstan, Karaganda, Kostanay, Pavlodar, North Kazakhstan, Turkestan, Abai, Zhetysu, Ulytau
Cluster 1 (low level, high growth rate)	Almaty, Astana, Shymkent, Mangystau, Kyzylorda

References

1. Quisumbing, A.; Cole, S.; Elias, M.; Faas, S.; Galiè, A.; Malapit, H.; Meinzen-Dick, R.; Myers, E.; Seymour, G.; Twyman, J. Measuring women's empowerment in agriculture: Innovations and evidence. *Glob. Food Secur.* **2023**, *38*, 100707. [CrossRef]
2. Meinzen-Dick, R.; Quisumbing, A.; Doss, C.; Theis, S. Women's land rights as a pathway to poverty reduction: Framework and review of available evidence. *Agric. Syst.* **2019**, *172*, 72–82. [CrossRef]
3. Diaz-Chavez, R. Socio-Economic Assessment of the Agriculture Sector and the Bioeconomy in East Africa—A Gender-Focused Approach. *Agriculture* **2025**, *15*, 1933. [CrossRef]
4. Wasif, M.; Sadiqe, M.; Khan, M.M.H.; Muzaffar, B. Challenges and Opportunities Faced by Rural Women Entrepreneurs. In *Examining Barriers and Building Resiliency for Rural Women Entrepreneurs*; IGI Global: Hershey, PA, USA, 2025; pp. 1–28. [CrossRef]
5. Namayengo, F.M.M.; van Ophem, J.A.C.; Antonides, G. A comparative study on the role of microcredit on agricultural production improvement among resource-poor rural women. *Front. Sustain. Food Syst.* **2023**, *7*, 1083660. [CrossRef]
6. Khan, C.M.; Kim, S.G. Evaluating the impact of agricultural credit access on smallholder maize farmers' productivity in the Northwest Region of Cameroon. *Sustainability* **2025**, *17*, 7574. [CrossRef]
7. Ullah, A.; Verner, V.; Madaki, M.Y.; Adams, F.; Bavorova, M. Factors influencing informal credit access and utilization among smallholder farmers: Insights from mountainous regions of Pakistan. *Agriculture* **2024**, *14*, 1764. [CrossRef]
8. Mbangiswano, S.Z.; Ndlovu, E.; Vuthela, Z.S. Empowering rural women agripreneurs through financial inclusion: Lessons from South Africa for the G20 development agenda. *Adm. Sci.* **2025**, *15*, 340. [CrossRef]
9. Yan, Z.; Wei, F.; Deng, X.; Li, C.; He, Q.; Qi, Y. Feminization of Agriculture: Do Female Farmers Have Higher Expectations for the Value of Their Farmland?—Empirical Evidence from China. *Agriculture* **2022**, *12*, 60. [CrossRef]
10. Lambrecht, I.B.; Mahrt, K.; Synt, N.L.K.; Win, H.E.; Win, K.Z. Gender gaps in land rights: Explaining different measures and why households differ in Myanmar. *Agric. Econ.* **2023**, *54*, 728–741. [CrossRef]
11. Baloi, V.A.; Belete, A.; Hlongwane, J.J. Women in agriculture: Irrigation as empowerment tool. *Am. J. Youth Women Empower.* **2023**, *2*, 1–8. [CrossRef]
12. Ragsdale, K.; Read-Wahidi, M.R.; Wei, T.; Martey, E.; Goldsmith, P. Using the WEAI+ to explore gender equity and agricultural empowerment: Baseline evidence among men and women smallholder farmers in Ghana's Northern Region. *J. Rural Stud.* **2018**, *64*, 123–134. [CrossRef]
13. Henrotte, S.; Van den Broeck, G. How does gender shape the adoption of sustainable agricultural practices? Evidence from Belgium. *J. Rural Stud.* **2026**, *121*, 103919. [CrossRef]
14. Asadullah, M.N.; Kambhampati, U. Feminization of farming, food security and female empowerment. *Glob. Food Secur.* **2021**, *29*, 100532. [CrossRef]
15. Ayamga, J.A.; Ayawine, A.; Ayentimi, D.T. Women empowerment and food-nutrition security in Sierra Leone: The Gender Model Family approach. *Agric. Food Secur.* **2023**, *12*, 44. [CrossRef]
16. Naheed, S.; Rukhsana, R. Transitioning to sustainable food systems in a changing climate and gender equality: A brief review. *Agric. Food Secur.* **2024**, *13*, 41. [CrossRef]
17. Timu, A.G.; Manoti, D.; Shee, A.; You, L. Impacts of gender-inclusive extension approaches on farmer understanding and willingness to pay for bundled financial services. *Curr. Res. Environ. Sustain.* **2024**, *8*, 100268. [CrossRef]

18. Agarwal, T.; Ravindranath, D.; Adamseged, M.E. *SoLAR Phase II Gender and Social Inclusion (GESI) Strategy*. Presented at the SoLAR Phase 2 Inception Meeting, New Delhi, India, 28 November 2025; International Water Management Institute: Colombo, Sri Lanka, 2025. Available online: https://hdl.handle.net/10568/180841 (accessed on 17 September 2026).
19. Azima, S.; Mundler, P. The gendered motives and experiences of Canadian women farmers in short food supply chains: Work satisfaction, values of care, and the potential for empowerment. *J. Rural Stud.* **2022**, *96*, 19–31. [CrossRef]
20. Rubin, D.; Boonabaana, B.; Manfre, C. Building an inclusive agriculture: Strengthening gender equality in agricultural value chains. In *2019 Annual Trends and Outlook Report: Gender Equality in Rural Africa: From Commitments to Outcomes*; Quisumbing, A.R., Meinzen-Dick, R.S., Njuki, J., Eds.; International Food Policy Research Institute: Washington, DC, USA, 2019; pp. 101–122. [CrossRef]
21. Ogunsola, O.E.; Adenuga, M.A.; Nnabueze, S.B. Fostering Inclusive Economies: The Role of Cooperatives in Empowering Women Entrepreneurs in Agriculture. *Glob. Multidiscip. Perspect. J.* **2024**, *1*, 26–46. [CrossRef]
22. Behera, S.K.; Panda, R.K.; Senapati, S. Role of financial and social capital in rural women micro-enterprises: Assessing entrepreneurial orientation as a performance catalyst. *J. Enterprising Communities People Places Glob. Econ.* **2026**, *20*, 59–93. [CrossRef]
23. Stawicka, E.; Parlinska, M. Female entrepreneurship in rural areas in the aspect of the labor market. In Proceedings of the International Scientific Conference on Economic Science for Rural Development, Jelgava, Latvia, 11–14 May 2021; Volume 55, pp. 396–404. [CrossRef]
24. Demedeme, G.; Opoku, C.B. Economic empowerment of rural women: Assessing the effectiveness of the rural enterprise program (REP) in Ghana, West Africa. *J. Agric. Ext. Rural Dev.* **2022**, *14*, 13–23. [CrossRef]
25. Mager, G.; Faße, A. The contribution of smallholders' livelihood activities on income inequality and poverty: Case study from rural Tanzania. *J. Int. Dev.* **2023**, *36*, 644–676. [CrossRef]
26. Ingutia, R.; Sumelius, J. Does cooperative membership facilitate access to credit for women farmers in rural Kenya? *J. Agric. Food Res.* **2024**, *18*, 101425. [CrossRef]
27. Nchanji, E.B.; Ndunguru, A.; Kabungo, C.; Katunzi, A.; Nyamolo, V.; Ouya, F.O.; Mutua, M.; Waswa, B.; Lutomia, C.K. Assessing gender disparities in farmers' access and use of climate-smart agriculture in Southern Tanzania. *Discov. Sustain.* **2025**, *6*, 337. [CrossRef]
28. Islam, M.M.; Jannat, A.; Al Rafi, D.A. Women participation in South Asian agriculture: A comprehensive systematic review. *Discov. Sustain.* **2024**, *5*, 490. [CrossRef]
29. Sassi, M. Women's empowerment in agriculture and household food security among smallholder farmers in Rejaf Payam, South Sudan: A mixed-methods analysis. *Agric. Food Econ.* **2026**, *14*, 35. [CrossRef]
30. Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan. Goal 5. Gender Equality. Available online: https://stat.gov.kz/ru/sustainable-development-goals/goal/5/ (accessed on 17 September 2026).
31. Wooldridge, J.M. Simple approaches to nonlinear difference-in-differences with panel data. *Econom. J.* **2023**, *26*, C31–C66. [CrossRef]
32. Chiu, A.; Lan, X.; Liu, Z.; Xu, Y. Causal panel analysis under parallel trends: Lessons from a large reanalysis study. *Am. Political Sci. Rev.* **2026**, *120*, 245–266. [CrossRef]
33. Lee, S.J.; Wooldridge, J.M. A Simple Transformation Approach to Difference-in-Differences Estimation for Panel Data. SSRN Working Paper 4516518. 2023. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4516518 (accessed on 17 September 2026). [CrossRef]
34. Tundui, H.P.; Tundui, C.S. Microcredit and women entrepreneurial success: A moderated mediation effect of household economic status. *Int. J. Sociol. Soc. Policy* **2024**, *44*, 793–808. [CrossRef]
35. Mazakaza, C.; Odoyo, C. The impact of micro financing on women: A brief review of the literature. *Int. J. Finance Account.* **2022**, *7*, 55–64. [CrossRef]
36. Shahriar, A.Z.M.; Chase, S.R.; Shepherd, D. Financial inclusion and low-income women's new venture initiation: A Field experiment. *J. Bus. Ventur.* **2025**, *40*, 106525. [CrossRef]
37. Bohn, S.; Wollni, M.; Paz, B. Cultivating Change: Exploring the Link Between Certification, Dietary Quality and Women's Empowerment Among Coffee Farmers in Rwanda. Sustainable Food Discussion Paper. 2024. Available online: https://hdl.handle.net/10419/301249 (accessed on 17 September 2026).
38. Mozumdar, L.; Lindgren, S.; Nishat, N. Modern agrotechnology, women's empowerment and poverty reduction nexus: Mediation of farm performance; empirical evidence on BAU-STR dryer. *World Dev. Perspect.* **2025**, *37*, 100673. [CrossRef]
39. Jemaneh, S.A.; Shibeshi, E.M. Women empowerment in agriculture and its effect on household food security: Evidence from Gamo Zone of Southern Ethiopia. *Agric. Food Secur.* **2023**, *12*, 37. [CrossRef]
40. Adefila, A.O.; Ajayi, O.O.; Toromade, A.S.; Sam-Bulya, N.J. A sociological review of gender equity in agricultural development: Global trends and lessons for US policy. *Int. J. Appl. Res. Soc. Sci.* **2024**, *6*, 2658–2677. [CrossRef]

41. Ge, Y.; Fan, L.; Li, Y.; Guo, J.; Niu, H. Gender differences in smallholder farmers' adoption of crop diversification: Evidence from Shaanxi Plain, China. *Clim. Risk Manag.* **2023**, *39*, 100482. [CrossRef]
42. Gil, J. Women's empowerment and crop diversification. *Nat. Food* **2023**, *4*, 639. [CrossRef]
43. Mohammed, F.Y.; Khonje, M.G.; Qaim, M. Women's roles in decision-making and nutrition-sensitive agriculture. *Food Secur.* **2025**, *17*, 1301–1315. [CrossRef]
44. Schoneveld, G.C.; Weng, X. Smallholder value creation in agrifood chains: Value network approach. *Land Use Policy* **2023**, *131*, 106676. [CrossRef]
45. Ghosh, S.; Neogi, C. Access to finance, entrepreneurship, and empowerment: A case study. In *Women's Entrepreneurship and Microfinance*; Springer: Singapore, 2017; pp. 173–189. [CrossRef]
46. Costa, V.; Piedrahita, N.; Mane, E.; Davis, B.; Slavchevska, V. Global estimates of women's and men's employment in agrifood systems from 2000 to 2021. *Glob. Food Secur.* **2026**, *48*, 100904. [CrossRef]
47. Hartarska, V.; Adjei, E.; Nadolnyak, D. The transition incentive program and women farmers in the USA. *Food Policy* **2024**, *129*, 102739. [CrossRef]
48. Amoah, J.O. Gendered challenges and coping strategies among smallholder farmers: An intersectionality analysis. *Soc. Sci. Humanit. Open* **2024**, *10*, 101028. [CrossRef]
49. Shali Lasrado, B.; Panakaje, N.; Parvin, S.R.; Kagzi, M.; Sheikh, N.; Irfana, S.; Banu, W. Skills assessment and development scale for rural women entrepreneurs: Conceptualization and validation. *J. Small Bus. Enterp. Dev.* **2025**, *32*, 1530–1557. [CrossRef]
50. Chance, E.A.; Abdoul, I.S. The roles and contributions of women to the health of their families and household economics in rural areas in the district of Mbe, Cameroon. *Soc. Sci. Humanit. Open* **2025**, *11*, 101456. [CrossRef]
51. Akybayeva, G.; Mussabekova, A.; Mambetova, S.; Ayaganova, M.; Koitanova, A. The impact of female entrepreneurship on economic growth in developing and developed economies. *Economics* **2024**, *12*, 145–162. [CrossRef]
52. Bekturganova, M.S.; Kurmasheva, M.T. Structural barriers and opportunities for financial inclusion of women in Kazakhstan. *Eurasian J. Econ. Bus. Stud.* **2025**, *69*, 19–33. [CrossRef]
53. Kuldasheva, Z.; Ahmad, M. Empowering economic growth through female labor force participation in Central Asia: Evidence from regression and dynamic analyses. *Asia Glob. Econ.* **2025**, *5*, 100115. [CrossRef]
54. Satpayeva, Z.T.; Kireyeva, A.A.; Kenzhegulova, G.; Yermekbayeva, D. Gender equality and women business of framework 5ms in Kazakhstan: Analysis and basic directions. *J. Asian Finance Econ. Bus.* **2020**, *7*, 253–263. [CrossRef]
55. Wang, S.; Mwalupaso, G.E.; Akter, A. Nexus of women's participation in smallholder financial inclusion programs and economic efficiency: Implications for poverty alleviation. *Front. Sustain. Food Syst.* **2025**, *9*, 1613678. [CrossRef]
56. Meuwissen, M.P.M.; Feindt, P.H.; Spiegel, A.; Termeer, C.J.A.M.; Mathijs, E.; de Mey, Y.; Finger, R.; Balmann, A.; Wauters, E.; Urquhart, J.; et al. A framework to assess the resilience of farming systems. *Agric. Syst.* **2019**, *176*, 102656. [CrossRef]
57. Lidder, P.; Cattaneo, A.; Chaya, M. Innovation and technology for achieving resilient and inclusive rural transformation. *Glob. Food Secur.* **2025**, *44*, 100827. [CrossRef]
58. Mundlak, Y. On the pooling of time series and cross section data. *Econometrica* **1978**, *46*, 69–85. [CrossRef]
59. Wooldridge, J.M. Two-way fixed effects, the two-way Mundlak regression, and difference-in-differences estimators. *Empir. Econ.* **2025**, *69*, 2545–2587. [CrossRef]
60. Cameron, A.C.; Gelbach, J.B.; Miller, D.L. Bootstrap-based improvements for inference with clustered errors. *Rev. Econ. Stat.* **2008**, *90*, 414–427. [CrossRef]
61. Roodman, D.; Nielsen, M.Ø.; MacKinnon, J.G.; Webb, M.D. Fast and wild: Bootstrap inference in Stata using boottest. *Stata J.* **2019**, *19*, 4–60. [CrossRef]
62. Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan. Gross Output of Agricultural Products (Services) by Region (Dynamic Series). Available online: https://stat.gov.kz/ru/industries/business-statistics/stat-forrest-village-hunt-fish/dynamic-tables/ (accessed on 7 October 2026).
63. Januzaj, Y.; Beqiri, E.; Luma, A. Determining the Optimal Number of Clusters using Silhouette Score as a Data Mining Technique. *Int. J. Online Biomed. Eng.* **2023**, *19*, 174–182. [CrossRef]
64. Hubert, L.; Arabie, P. Comparing partitions. *J. Classif.* **1985**, *2*, 193–218. [CrossRef]
65. Hennig, C. Cluster-wise assessment of cluster stability. *Comput. Stat. Data Anal.* **2007**, *52*, 258–271. [CrossRef]
66. Salman, H.A.; Kalakech, A.; Steiti, A. Random forest algorithm overview. *Babylon. J. Mach. Learn.* **2024**, *2024*, 69–79. [CrossRef]
67. Mohapatra, N.; Shreya, K.; Chinmay, A. Optimization of the random forest algorithm. In *Advances in Data Science and Management: Proceedings of ICDSM 2019*; Springer: Singapore, 2020; pp. 201–208. [CrossRef]
68. Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan. Number of Active Small and Medium-Sized Enterprises by Region. Available online: https://gender.stat.gov.kz/ru/category/6 (accessed on 7 October 2026).

69. AbdulKareem, H.K.; Omenka, L.C.O.; Assoua, J.E.; Okunade, S.O.; Bomdzele, E., Jr.; Omosehin, O.O.; Binuomote, S.O.; Enwa, S.; Osabuohien, E.; Akpan, U.S. Sustainable agribusiness, food security, and women's empowerment: Rural women at the heart of the virtuous cycle. In *Agripreneurship, Digital Inclusion and Sustainability: Rural Innovation and Growth in the Global South*; CABI: Wallingford, UK, 2026; pp. 32–50. [CrossRef]
70. Naika, R.; Pawar, S. Women-led and rural sustainable agricultural livelihoods pathways to achieving sustainable development goals by 2030. *Discov. Glob. Soc.* **2026**, *4*, 3. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.